Evaluation of PREEMPT_RT in virtualized environments
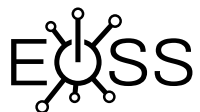
Jan Altenberg
Open Source Automation Development Lab (OSADL) eG

Agenda

- Software containerization
  - Docker container

- Hardware virtualization
  - KVM
  - Jailhouse

- Impact of shared resources

To be evaluated

- Real-time behavior on the host

- Real-time behavior on the guest

- Level of isolation

EMBEDDED
OPEN SOURCE
SUMMIT

Evaluation of PREEMPT_RT in virtualized environments
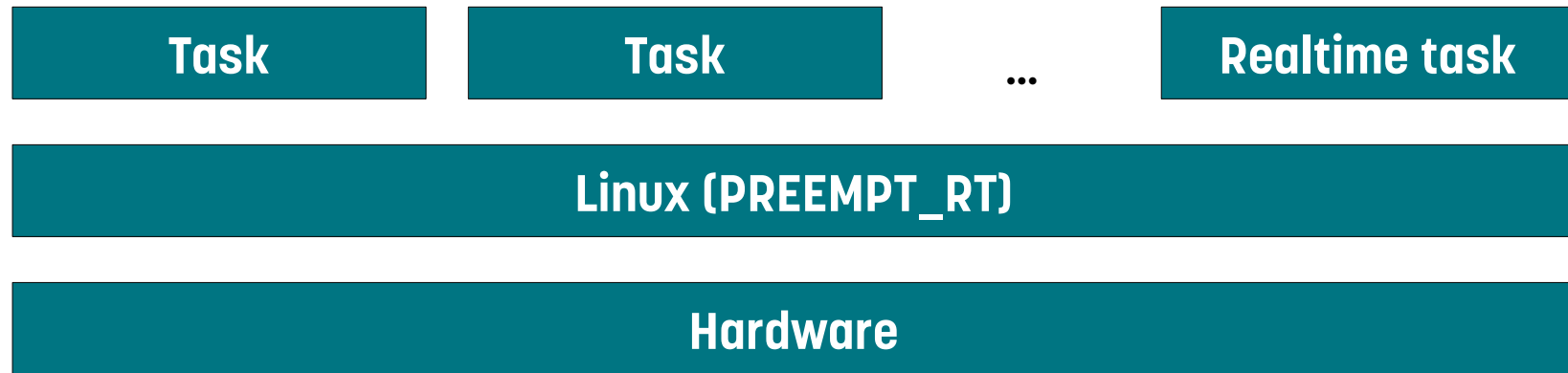Embedded Open Source Summit
June 29, 2023, Prague, Czech Republic

OSADL
Open Source Automation Development Lab eG

PREEMPT_RT and containerization

Task	Task	...	Realtime task

Linux (PREEMPT_RT)

Hardware

PREEMPT_RT and containerization

cyclictest

cyclictest

Docker container (Rack B Slot 2s)

Host (Rack 0 Slot 2)

Hardware

Latency measurement

cyclictest
Master process

T_1

T_2

T_N

cyclictest
Meas. thread

cyclictest
Meas. thread

cyclictest
Meas. thread

Evaluation of PREEMPT_RT in virtualized environments
Embedded Open Source Summit
June 29, 2023, Prague, Czech Republic

Latency measurement

Run on CPU #0

Priority 99

High-resolution timer

Start one thread

Interval 100 µs

```
# cyclictest -n -a0 -t1  -p99 -i100

T: 0 (26623) P:99 I:100 C:2244063 Min:  2 Act:  5 Avg:  4 Max:   27
```

Process ID

Cycles

Worst-case latency in µs

EMBEDDED
OPEN SOURCE
SUMMIT

OSADL
Open Source Automation Development Lab eG

Always remember the general principles

- Real-time behavior can only be guaranteed, if the real-time task has – **during its entire run time** – the **highest priority** of all tasks running on the same core.

- A real-time task is preferably **restricted to run on a particular core** by setting the affinity mask accordingly.

- These general principles are also valid when running in containerized or virtualized environments.

EMBEDDED
OPEN SOURCE
SUMMIT

Evaluation of PREEMPT_RT in virtualized environments
Embedded Open Source Summit
June 29, 2023, Prague, Czech Republic

OSADL
Open Source Automation Development Lab eG

Latency measurement

```
# cyclictest -n -a0 -t12 -p99 -i100 -d0
T: 0 ( 2910) P:99 I:100 C:3217008 Min:   2 Act:   6 Avg:   4 Max:     32
T: 1 ( 2911) P:98 I:100 C:3217008 Min:   1 Act:   4 Avg:   3 Max:     59
T: 2 ( 2912) P:97 I:100 C:3217007 Min:   2 Act:   4 Avg:   3 Max:     47
T: 3 ( 2913) P:96 I:100 C:3217007 Min:   2 Act:  11 Avg:   3 Max:     53
T: 4 ( 2914) P:95 I:100 C:3217007 Min:   2 Act:   9 Avg:   4 Max:     53
T: 5 ( 2915) P:94 I:100 C:3217007 Min:   3 Act:   9 Avg:   7 Max:     89
T: 6 ( 2916) P:93 I:100 C:3217007 Min:   2 Act:   5 Avg:   4 Max:     85
T: 7 ( 2917) P:92 I:100 C:3217006 Min:   2 Act:  10 Avg:   5 Max:    119
T: 8 ( 2918) P:91 I:100 C:3217006 Min:   2 Act:  13 Avg:   9 Max:    148
T: 9 ( 2919) P:90 I:100 C:3217007 Min:   1 Act:   4 Avg:   4 Max:    178
T:10 ( 2920) P:89 I:100 C:3217006 Min:   1 Act:   4 Avg:   3 Max:   1413
T:11 ( 2921) P:88 I:100 C:3217006 Min:   3 Act:   7 Avg:  10 Max:  27331
```

EMBEDDED
OPEN SOURCE
SUMMIT

Evaluation of PREEMPT_RT in virtualized environments
Embedded Open Source Summit
June 29, 2023, Prague, Czech Republic

OSADL
Open Source Automation Development Lab eG

Latency measurement

```
# cyclictest -n -a0 -t12 -p99 -i100 -d0
T: 0 ( 2910) P:99 I:100 C:3217008 Min:   2 Act:  6 Avg:  4 Max:    32
T: 1 ( 2911) P:98 I:100 C:3217008 Min:   1 Act:  4 Avg:  3 Max:    59
T: 2 ( 2912) P:97 I:100 C:3217007 Min:   2 Act:  4 Avg:  3 Max:    47
T: 3 ( 2913) P:96 I:100 C:3217007 Min:   2 Act: 11 Avg:  3 Max:    53
T: 4 ( 2914) P:95 I:100 C:3217007 Min:   2 Act:  9 Avg:  4 Max:    53
T: 5 ( 2915) P:94 I:100 C:3217007 Min:   3 Act:  9 Avg:  7 Max:    89
T: 6 ( 2916) P:93 I:100 C:3217007 Min:   2 Act:  5 Avg:  4 Max:    85
T: 7 ( 2917) P:92 I:100 C:3217006 Min:   2 Act: 10 Avg:  5 Max:   119
T: 8 ( 2918) P:91 I:100 C:3217006 Min:   2 Act: 13 Avg:  9 Max:   148
T: 9 ( 2919) P:90 I:100 C:3217007 Min:   1 Act:  4 Avg:  4 Max:   178
T:10 ( 2920) P:89 I:100 C:3217006 Min:   1 Act:  4 Avg:  3 Max:  1413
T:11 ( 2921) P:88 I:100 C:3217006 Min:   3 Act:  7 Avg: 10 Max: 27331
```
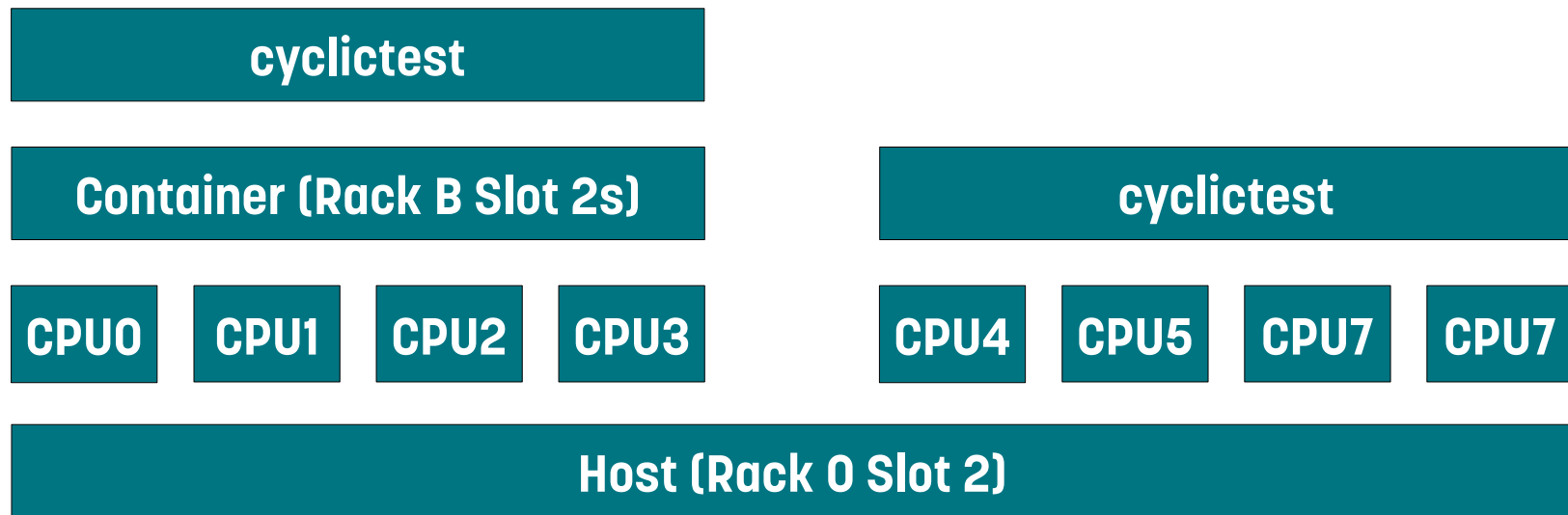
THE SYSTEM IS OVERCOMMITTED

EMBEDDED OPEN SOURCE SUMMIT

OSADL
Open Source Automation Development Lab eG

Latency measurement

```
# cyclictest        -S      -p99 -i100 -d0
T: 0 (15350) P:99 I:100 C:3839755 Min:   2 Act:   6 Avg:   3 Max:     24
T: 1 (15351) P:99 I:100 C:3839755 Min:   2 Act:   7 Avg:   4 Max:     19
T: 2 (15352) P:99 I:100 C:3839755 Min:   2 Act:   8 Avg:   4 Max:     27
T: 3 (15353) P:99 I:100 C:3839755 Min:   2 Act:   5 Avg:   4 Max:     24
T: 4 (15354) P:99 I:100 C:3839755 Min:   2 Act:   5 Avg:   3 Max:     20
T: 5 (15355) P:99 I:100 C:3839755 Min:   2 Act:   5 Avg:   5 Max:     52
T: 6 (15356) P:99 I:100 C:3839755 Min:   2 Act:   5 Avg:   4 Max:     20
T: 7 (15357) P:99 I:100 C:3839755 Min:   2 Act:   5 Avg:   3 Max:     17
T: 8 (15358) P:99 I:100 C:3839755 Min:   2 Act:  10 Avg:   4 Max:     28
T: 9 (15359) P:99 I:100 C:3839754 Min:   2 Act:   5 Avg:   4 Max:     22
T:10 (15360) P:99 I:100 C:3839755 Min:   2 Act:   5 Avg:   4 Max:     42
T:11 (15361) P:99 I:100 C:3839755 Min:   2 Act:   5 Avg:   5 Max:     34
```

EMBEDDED
OPEN SOURCE
SUMMIT

Evaluation of PREEMPT_RT in virtualized environments
Embedded Open Source Summit
June 29, 2023, Prague, Czech Republic

OSADL
Open Source Automation Development Lab eG

Partitioning of the system resources

cyclictest

Container (Rack B Slot 2s)		cyclictest

CPU0	CPU1	CPU2	CPU3		CPU4	CPU5	CPU7	CPU7

Host (Rack 0 Slot 2)

Partitioning of the system resources

cyclictest -a0-3 -t4

cyclictest

Container (Rack B Slot 2s)					cyclictest			
CPU0	CPU1	CPU2	CPU3		CPU4	CPU5	CPU7	CPU7

Host (Rack O Slot 2)

EMBEDDED
OPEN SOURCE
SUMMIT

Evaluation of PREEMPT_RT in virtualized environments
Embedded Open Source Summit
June 29, 2023, Prague, Czech Republic

OSADL
Open Source Automation Development Lab eG

Partitioning of the system resources

cyclictest **-a0-3** -t4

cyclictest **-a4-7** -t4

cyclictest

Container (Rack B Slot 2s)

cyclictest

CPU0　**CPU1**　**CPU2**　**CPU3**

CPU4　**CPU5**　**CPU7**　**CPU7**

Host (Rack O Slot 2)

EMBEDDED OPEN SOURCE SUMMIT

OSADL
Open Source Automation Development Lab eG

Latency measurement on the host (load)

Evaluation of PREEMPT_RT in virtualized environments
Embedded Open Source Summit
June 29, 2023, Prague, Czech Republic
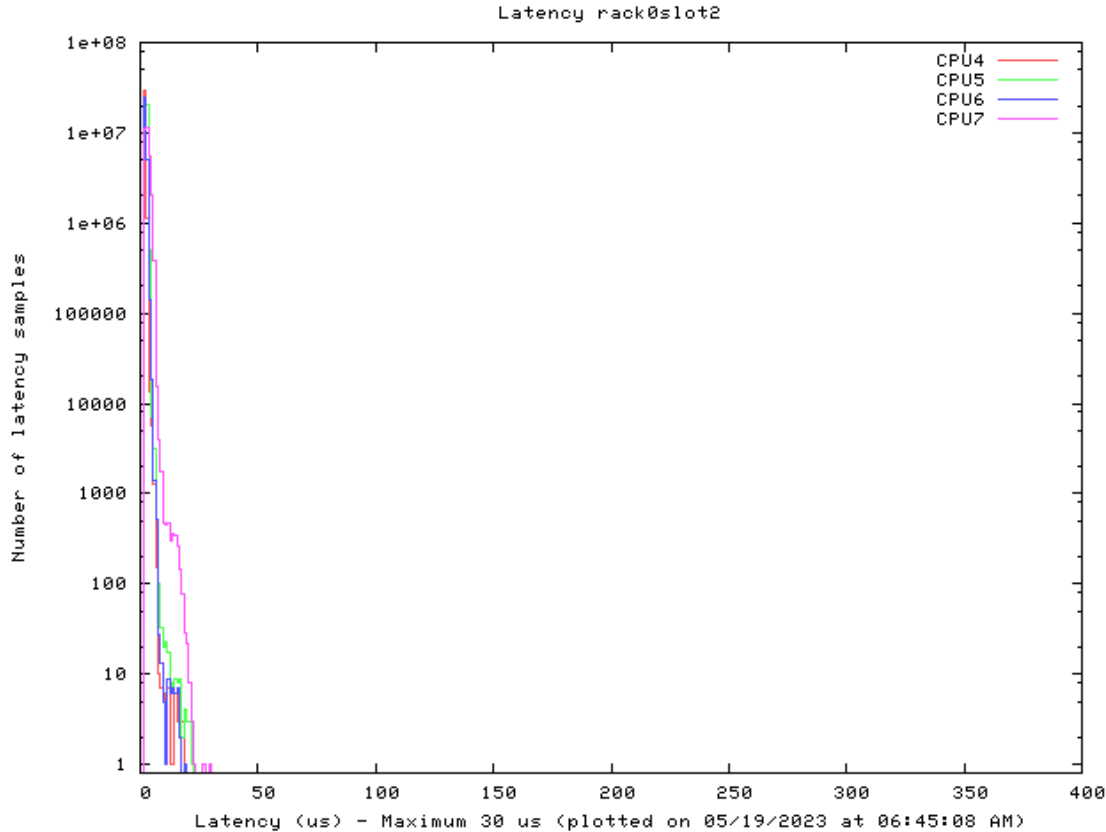
Latency measurement in the container (load)

Latency rackbslot2s

CPU0
CPU1
CPU2
CPU3

Number of latency samples

Latency (us) - Maximum 26 us (plotted on 05/18/23 at 18:45:02)

Evaluation of PREEMPT_RT in virtualized environments
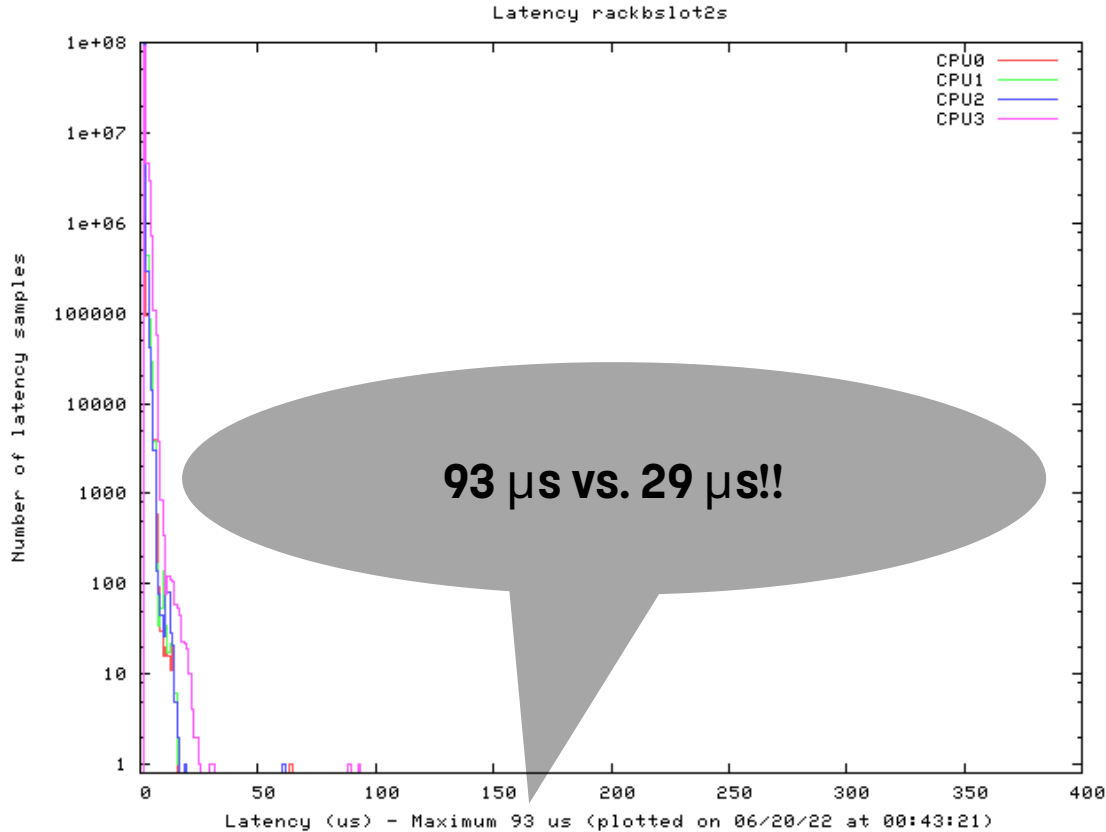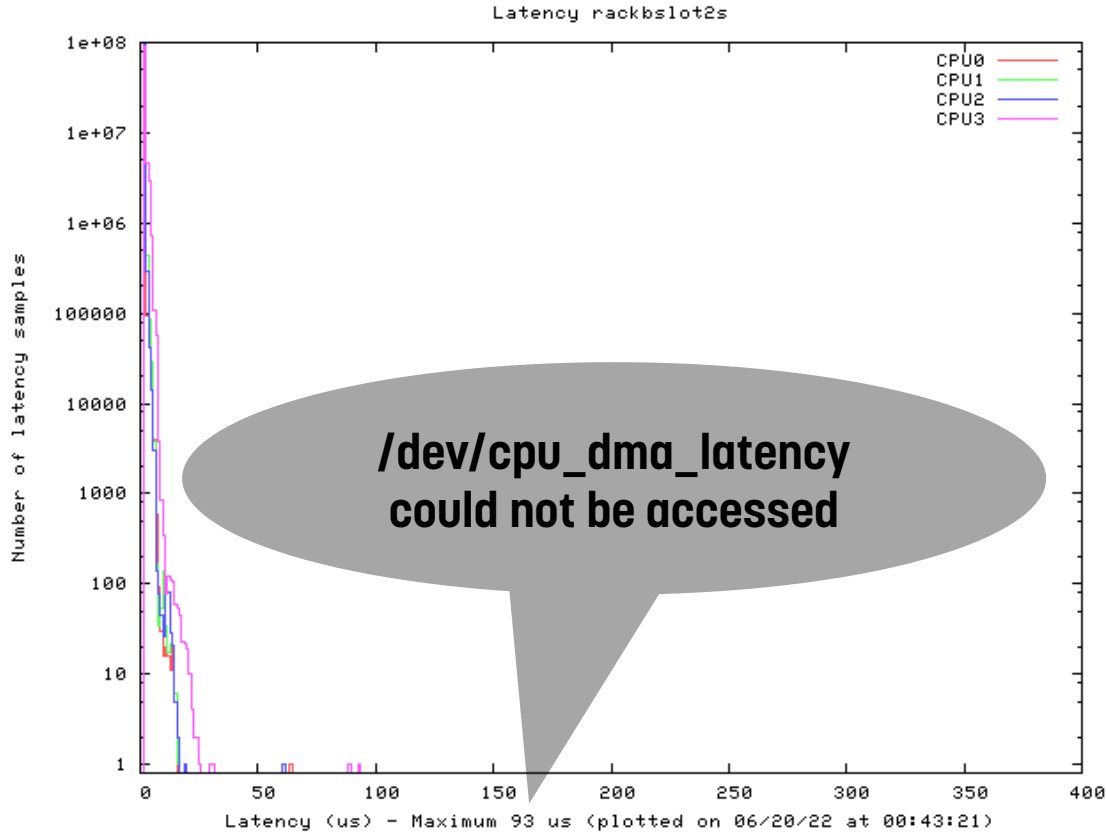Embedded Open Source Summit
June 29, 2023, Prague, Czech Republic
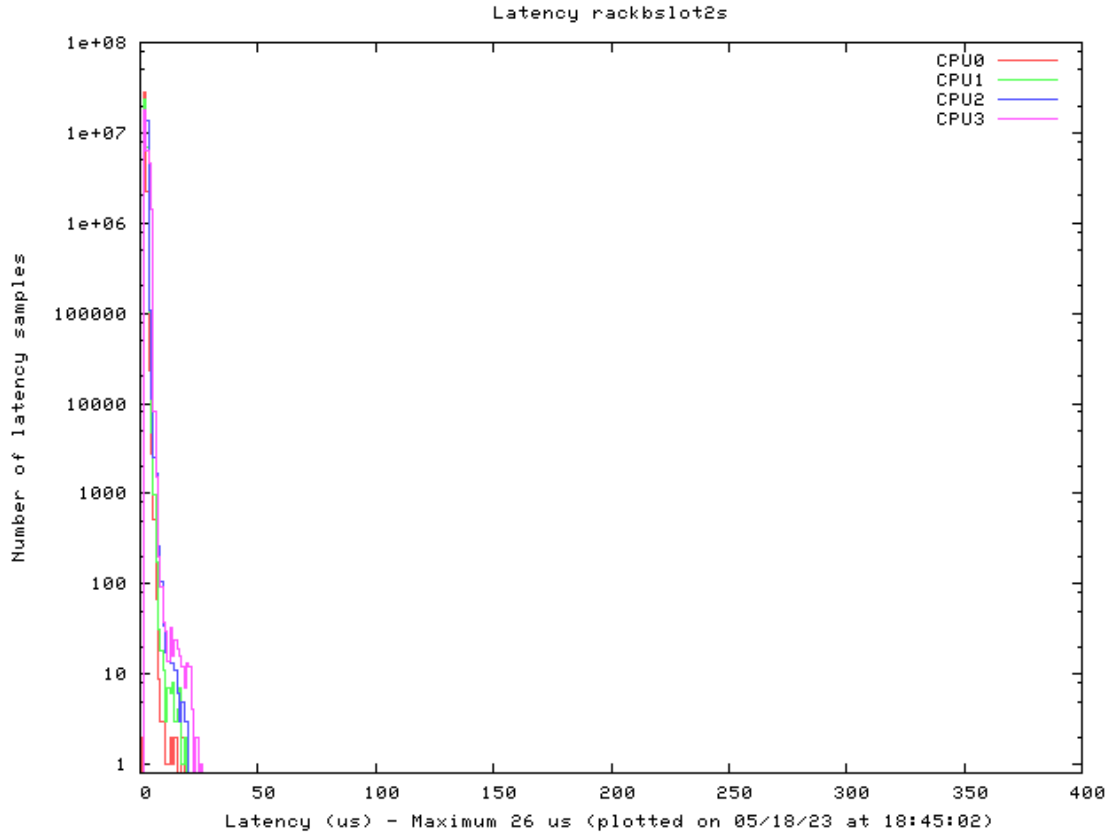
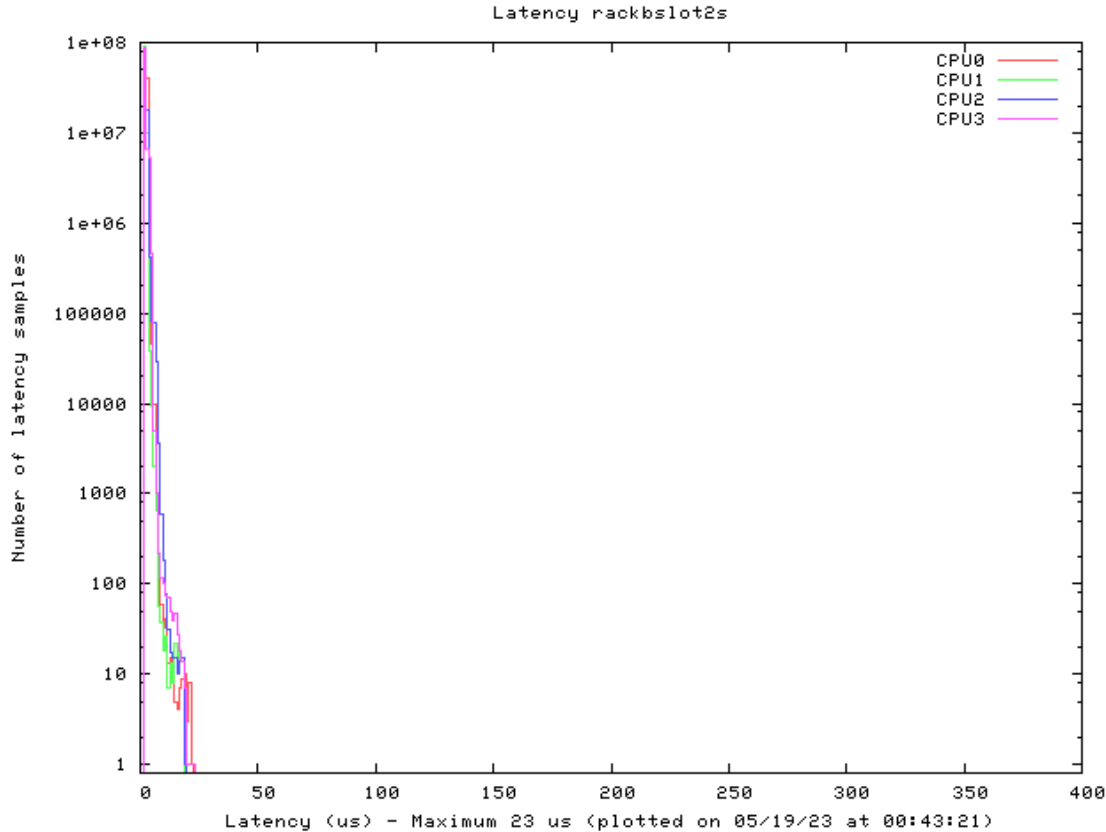# Latency measurement on the host (idle)
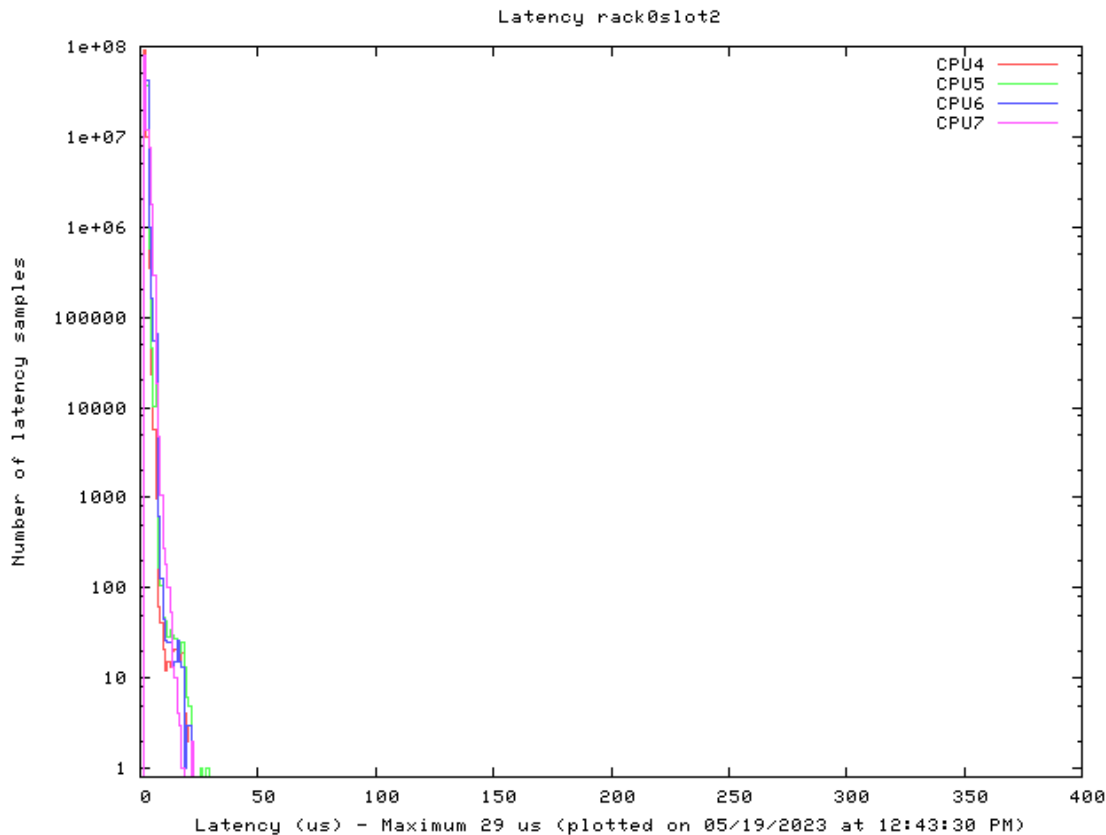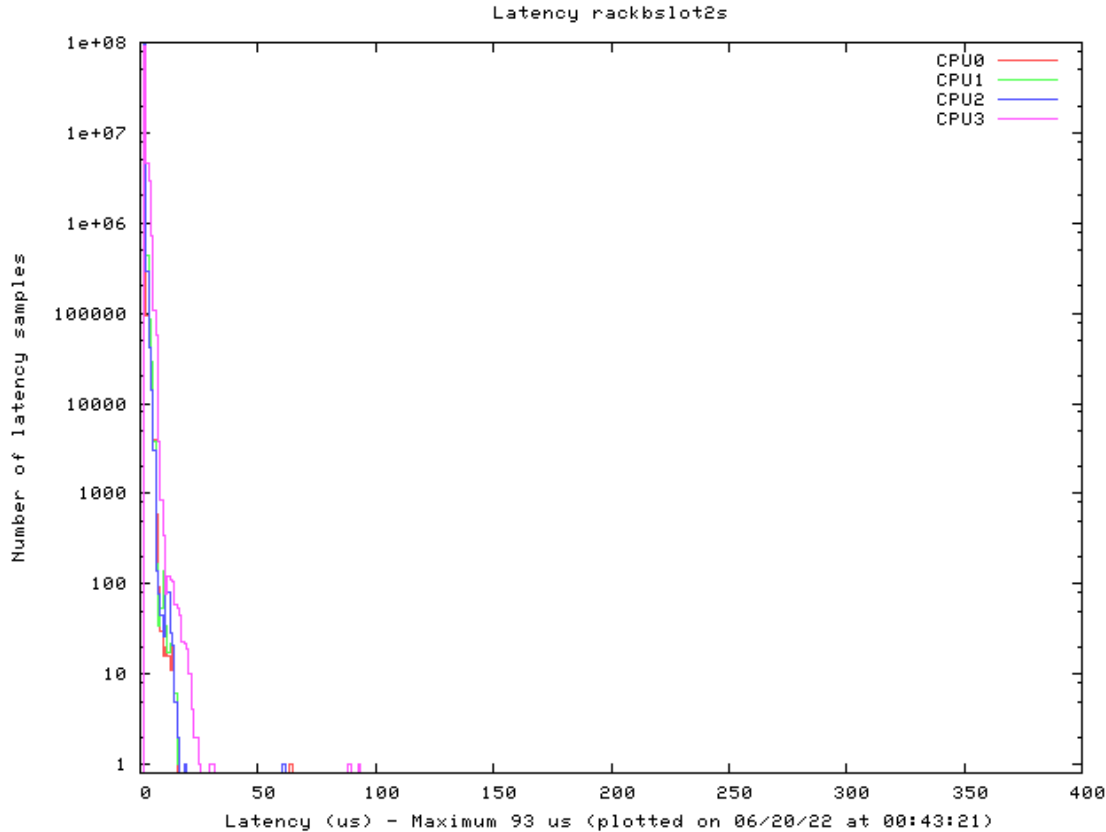
Latency measurement in the container (idle)

Latency measurement in the container (idle)

93 µs vs. 29 µs!!

Evaluation of PREEMPT_RT in virtualized environments
Embedded Open Source Summit
June 29, 2023, Prague, Czech Republic

Latency measurement in the container (idle)

Evaluation of PREEMPT_RT in virtualized environments
Embedded Open Source Summit
June 29, 2023, Prague, Czech Republic

Latency measurement in the container (idle)

Evaluation of PREEMPT_RT in virtualized environments
Embedded Open Source Summit
June 29, 2023, Prague, Czech Republic

Blocksys

- Blocksys is a Linux **kernel module**
- Blocks the system for a specified number of cycles
    - Preemption disabled
    - Local IRQ processing disabled
- This can be accessed as device over the virtual filesystem in **/dev/blocksys**
- Can be used to artificially create a system latency

Blocksys on docker host, measurement on host

```
[root@rack0slot2 ~]# /usr/local/bin/cyclictest -m -n -Sp99 -i 200 -d0
# /dev/cpu_dma_latency set to 0us
policy: fifo: loadavg: 3.28 1.93 1.04 1/719 4162354

T: 0 (4162327) P:99 I:200 C:  92855 Min:       4 Act:      5 Avg:      5 Max:      14
T: 1 (4162328) P:99 I:200 C:  92771 Min:       5 Act:      5 Avg:      5 Max:      18
T: 2 (4162329) P:99 I:200 C:  92815 Min:       5 Act:      5 Avg:      5 Max:      16
T: 3 (4162330) P:99 I:200 C:  92794 Min:       4 Act:      5 Avg:      5 Max:      16
T: 4 (4162331) P:99 I:200 C:  92774 Min:       5 Act:      5 Avg:      5 Max:      18
T: 5 (4162332) P:99 I:200 C:  92753 Min:       5 Act:      6 Avg:      5 Max:      18
T: 6 (4162333) P:99 I:200 C:  92731 Min:       5 Act:      5 Avg:      5 Max:      16
T: 7 (4162334) P:99 I:200 C:  92710 Min:       5 Act:      6 Avg:      5 Max:      18
```

EMBEDDED
OPEN SOURCE
SUMMIT

OSADL
Open Source Automation Development Lab eG

Blocksys on docker host, measurement on host

```
[root@rack0slot2 ~]# /usr/local/bi      clictest -m -n -Sp99 -i 200 -d0
# /dev/cpu_dma_latency set to 0u
policy: fifo: loadavg: 3.28 1.9    04 1/719 4162354

T: 0 (4162327) P:99 I:200 C:    855            4 Act:    5 Avg:    5 Max:      14
T: 1 (4162328) P:99 I:200             blocksys   5 Act:    5 Avg:    5 Max:      18
T: 2 (4162329) P:99 I:200 C:  92815 Mi          5 Act:    5 Avg:    5 Max:      16
T: 3 (4162330) P:99 I:200 C:  92794             4 Act:    5 Avg:    5 Max:      16
T: 4 (4162331) P:99 I:200 C:  9277  n:          5 Act:    5 Avg:    5 Max:      18
T: 5 (4162332) P:99 I:200 C:  927   Min:        5 Act:    6 Avg:    5 Max:      18
T: 6 (4162333) P:99 I:200     9     Min:        5 Act:    5 Avg:    5 Max:      16
T: 7 (4162334) P:99 I:200      10 Min:          5 Act:    6 Avg:    5 Max:      18
```

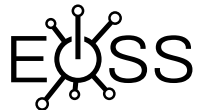# Blocksys on docker host, measurement on host

```
[root@rack0slot2 ~]# /usr/local/bin/cyclictest -m -n -Sp99 -i 200 -d0
# /dev/cpu_dma_latency set to 0us
policy: fifo: loadavg: 1.66 1.68 1.03 1/718 4165013

T: 0 (4164963) P:99 I:200 C: 105863 Min:        5 Act:     5 Avg:     5 Max:      4041
T: 1 (4164964) P:99 I:200 C: 105853 Min:        5 Act:     5 Avg:     5 Max:      4063
T: 2 (4164965) P:99 I:200 C: 105832 Min:        5 Act:     5 Avg:     5 Max:      4099
T: 3 (4164966) P:99 I:200 C: 105778 Min:        5 Act:     6 Avg:     5 Max:      4200
T: 4 (4164967) P:99 I:200 C: 105792 Min:        5 Act:     5 Avg:     5 Max:      4031
T: 5 (4164968) P:99 I:200 C: 105772 Min:        5 Act:     6 Avg:     5 Max:      4038
T: 6 (4164969) P:99 I:200 C: 105751 Min:        5 Act:     5 Avg:     5 Max:      4000
T: 7 (4164970) P:99 I:200 C: 105730 Min:        5 Act:     5 Avg:     5 Max:      4142
```

EMBEDDED OPEN SOURCE SUMMIT

Evaluation of PREEMPT_RT in virtualized environments
Embedded Open Source Summit
June 29, 2023, Prague, Czech Republic

OSADL
Open Source Automation Development Lab eG

Blocksys (docker host), measurement on container

```
root@rackbslot2s.osadl.org:~# /usr/local/bin/cyclictest -m -n -Sp99 -i 200 -d0
policy: fifo: loadavg: 0.93 0.58 0.68 2/716 4170637

T: 0 (4170619) P:99 I:200 C:   43486 Min:      4 Act:    5 Avg:    5 Max:      16
T: 1 (4170620) P:99 I:200 C:   43473 Min:      4 Act:    5 Avg:    5 Max:      14
T: 2 (4170621) P:99 I:200 C:   43458 Min:      4 Act:    5 Avg:    5 Max:      18
T: 3 (4170622) P:99 I:200 C:   43444 Min:      4 Act:    6 Avg:    5 Max:      22
T: 4 (4170623) P:99 I:200 C:   43429 Min:      4 Act:    5 Avg:    5 Max:      18
T: 5 (4170624) P:99 I:200 C:   43414 Min:      4 Act:    6 Avg:    5 Max:      13
T: 6 (4170625) P:99 I:200 C:   43399 Min:      4 Act:    6 Avg:    5 Max:      14
T: 7 (4170626) P:99 I:200 C:   43384 Min:      5 Act:    8 Avg:    6 Max:      19
```
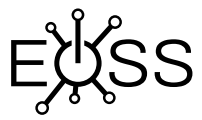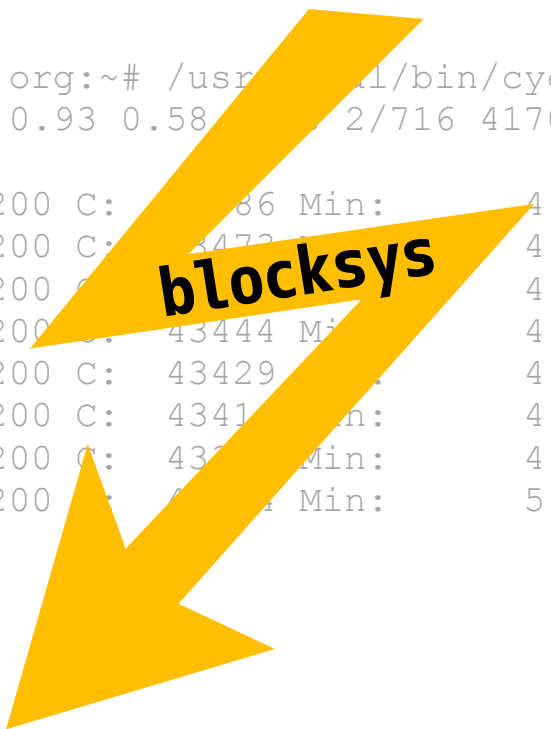
EMBEDDED OPEN SOURCE SUMMIT

OSADL
Open Source Automation Development Lab eG

Blocksys (docker host), measurement on container

```
root@rackbslot2s.osadl.org:~# /usr     l/bin/cyclictest -m -n -Sp99 -i 200 -d0
policy: fifo: loadavg: 0.93 0.58      2/716 4170637

T: 0 (4170619) P:99 I:200 C:    86 Min:     4 Act:    5 Avg:    5 Max:    16
T: 1 (4170620) P:99 I:200 C:  472  4 Act:    5 Avg:    5 Max:    14
T: 2 (4170621) P:99 I:200      4 Act:    5 Avg:    5 Max:    18
T: 3 (4170622) P:99 I:200     43444 Min:   4 Act:    6 Avg:    5 Max:    22
T: 4 (4170623) P:99 I:200 C:  43429      4 Act:    5 Avg:    5 Max:    18
T: 5 (4170624) P:99 I:200 C:  4341   n:   4 Act:    6 Avg:    5 Max:    13
T: 6 (4170625) P:99 I:200    43   Min:   4 Act:    6 Avg:    5 Max:    14
T: 7 (4170626) P:99 I:200          Min:   5 Act:    8 Avg:    6 Max:    19
```

blocksys

Blocksys (docker host), measurement on container

```
root@rackbslot2s.osadl.org:~# /usr/local/bin/cyclictest -m -n -Sp99 -i 200 -d0
policy: fifo: loadavg: 2.01 0.85 0.77 1/717 4170690

T: 0 (4170643) P:99 I:200 C:   90079 Min:        4 Act:    7 Avg:    5 Max:    3678
T: 1 (4170644) P:99 I:200 C:   90064 Min:        4 Act:    5 Avg:    5 Max:    3767
T: 2 (4170645) P:99 I:200 C:   90051 Min:        5 Act:    6 Avg:    5 Max:    3721
T: 3 (4170646) P:99 I:200 C:   90035 Min:        4 Act:    5 Avg:    5 Max:    3808
T: 4 (4170647) P:99 I:200 C:   90022 Min:        4 Act:    5 Avg:    5 Max:    3647
T: 5 (4170648) P:99 I:200 C:   90007 Min:        4 Act:    5 Avg:    5 Max:    3825
T: 6 (4170649) P:99 I:200 C:   89979 Min:        4 Act:    6 Avg:    5 Max:    3775
T: 7 (4170650) P:99 I:200 C:   89979 Min:        4 Act:    5 Avg:    5 Max:    3807
```
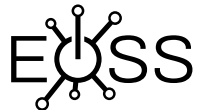
EMBEDDED
OPEN SOURCE
SUMMIT

OSADL
Open Source Automation Development Lab eG

Setup of the system for KVM tests

- Hardware: Intel Core i7 processor with 6 physical cores

Host system with PREEMPT_RT Linux

Guest system
- KVM (Intel VT-x) with PREEMPT_RT Linux

| Core #0 | Core #1 | Core #2 | | Core #3 | Core #4 | Core #5 |

Setup of the system for KVM tests

Isolating cores #3, #4 and #5 from the host system:

```
isolcpus=3,4,5
rcu_nocbs=3,4,5
nohz_full=3,4,5
irqaffinity=0,1,2
```

Host system with PREEMPT_RT Linux

Guest system:
- KVM (Intel VT-x) with PREEMPT_RT Linux

| Core #0 | Core #1 | Core #2 |

| Core #3 | Core #4 | Core #5 |

Setup of the system for KVM tests

Pinning the virtual CPUs to the isolated cores:

```
<cputune>
  <vcpupin vcpu='0' cpuset='3'/>
  <vcpupin vcpu='1' cpuset='4'/>
  <vcpupin vcpu='2' cpuset='5'/>
  <vcpusched vcpus='0-2' scheduler='fifo' priority='1'/>
</cputune>
```
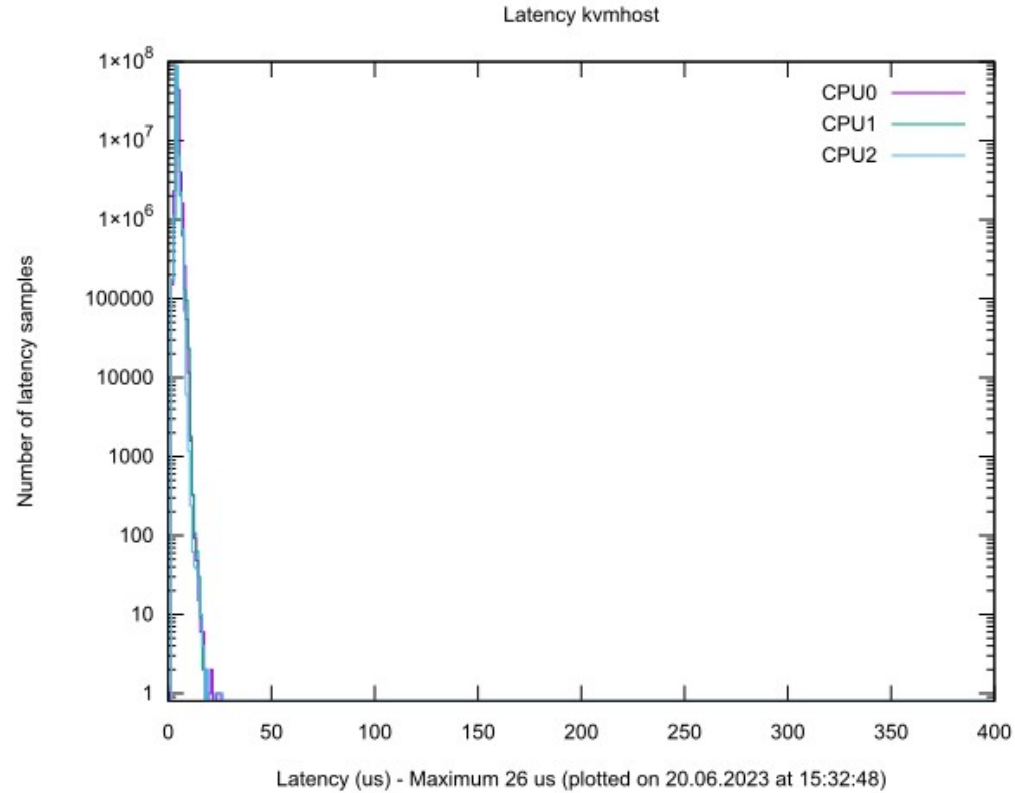
Guest system
• KVM (Intel VT-x) with PREEMPT_RT Linux

Host system with PREEMPT_RT Linux

| Core #0 | Core #1 | Core #2 | | Core #3 | Core #4 | Core #5 |

EMBEDDED
OPEN SOURCE
SUMMIT

Evaluation of PREEMPT_RT in virtualized environments
Embedded Open Source Summit
June 29, 2023, Prague, Czech Republic

OSADL
Open Source Automation Development Lab eG

Intel VT-x via KVM

Host latency: 26 μs

Evaluation of PREEMPT_RT in virtualized environments
Embedded Open Source Summit
June 29, 2023, Prague, Czech Republic

Intel VT-x via KVM (1)

Guest latency: 606 µs
More than 20 times longer than host latency

Latency kvmguest

Number of latency samples vs. Latency (us) - Maximum 606 us (plotted on 20.06.2023 at 15:30:59)

CPU0, CPU1, CPU2

EMBEDDED OPEN SOURCE SUMMIT

OSADL
Open Source Automation Development Lab eG
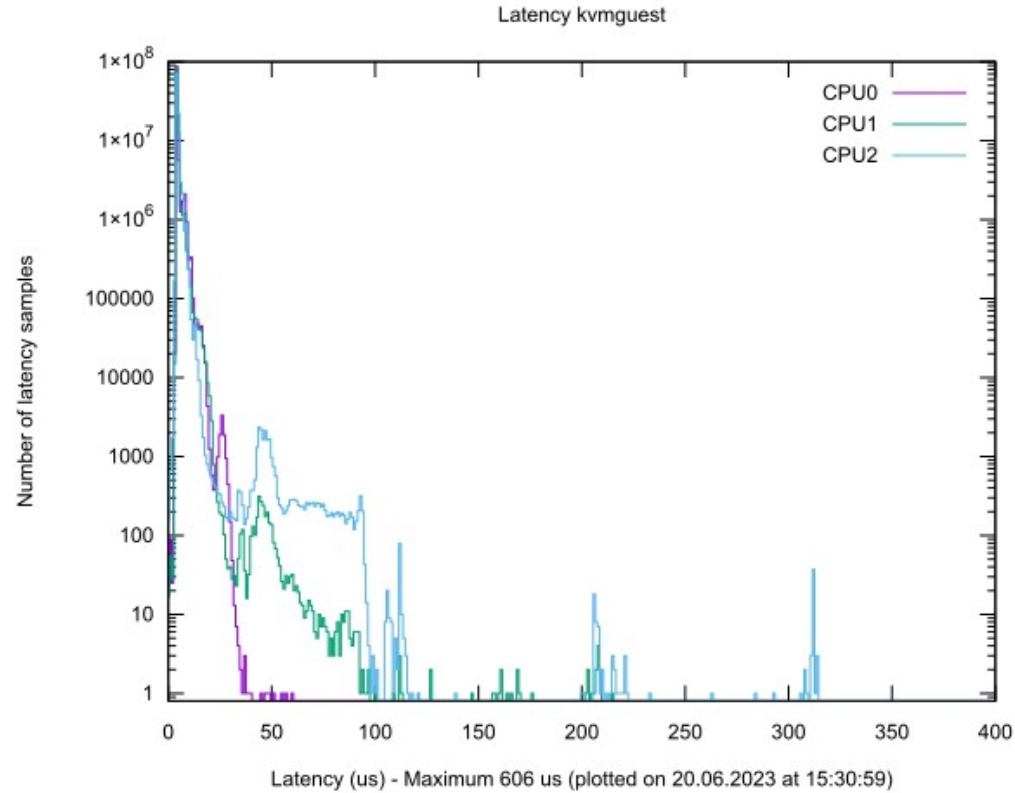
Intel VT-x via KVM (2)

Guest latency: 448 µs

Evaluation of PREEMPT_RT in virtualized environments
Embedded Open Source Summit
June 29, 2023, Prague, Czech Republic

KVM: Blocksys on the host system

Host system with PREEMPT_RT Linux

blocksys

Core #0   Core #1   Core #2

Guest system
- KVM (Intel VT-x) with PREEMPT_RT Linux

Core #3   Core #4   Core #5

EMBEDDED OPEN SOURCE SUMMIT

Evaluation of PREEMPT_RT in virtualized environments
Embedded Open Source Summit
June 29, 2023, Prague, Czech Republic

OSADL
Open Source Automation Development Lab eG

KVM: Blocksys on the host system

```
root@kvmrt:~# cyclictest -m -t3 -a 0-2 -p99 -i200
# /dev/cpu_dma_latency set to 0us
policy: fifo: loadavg: 4.89 2.51 1.54 1/841 5062

T: 0 ( 5043) P:99 I:200 C: 356988 Min:      2 Act:    4 Avg:    4 Max:    4018
T: 1 ( 5044) P:99 I:200 C: 356988 Min:      2 Act:    4 Avg:    4 Max:    3812
T: 2 ( 5045) P:99 I:200 C: 356875 Min:      2 Act:    3 Avg:    3 Max:    4043


root@kvmrtguest:~# cyclictest -m -Sp99 -i200
# /dev/cpu_dma_latency set to 0us
policy: fifo: loadavg: 2.89 2.23 1.19 1/267 767

T: 0 (  765) P:99 I:200 C:  74813 Min:      4 Act:   13 Avg:   12 Max:    4055
T: 1 (  766) P:99 I:200 C:  74786 Min:      4 Act:   10 Avg:   11 Max:    3917
T: 2 (  767) P:99 I:200 C:  74777 Min:      4 Act:   11 Avg:   12 Max:    4086
```
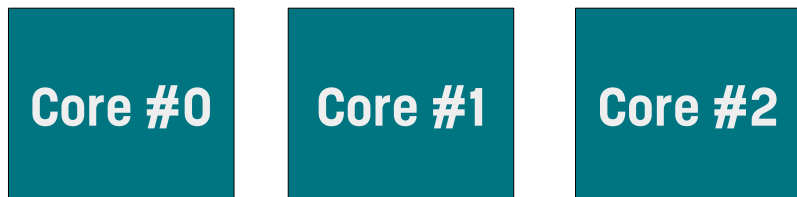
EMBEDDED OPEN SOURCE SUMMIT

OSADL
Open Source Automation Development Lab eG

KVM: Blocksys in the guest system

Host system with PREEMPT_RT Linux

Guest system
• KVM (Intel ...) with PREEMPT ... Linux

blocksys

| Core #0 | Core #1 | Core #2 | Core #3 | Co... #4 | Core #5 |

KVM: Blocksys in the guest system

```
root@kvmrt:~# cyclictest -m -t3 -a 0-2 -p99 -i200 -d0
# /dev/cpu_dma_latency set to 0us
policy: fifo: loadavg: 2.11 2.07 1.50 2/845 5090

T: 0 ( 5088) P:99 I:200 C:  32077 Min:      3 Act:    4 Avg:    3 Max:      8
T: 1 ( 5089) P:99 I:200 C:  32077 Min:      3 Act:    4 Avg:    4 Max:      9
T: 2 ( 5090) P:99 I:200 C:  32076 Min:      2 Act:    3 Avg:    3 Max:     10


root@kvmrtguest:~# cyclictest -m -Sp99 -i200
# /dev/cpu_dma_latency set to 0us
policy: fifo: loadavg: 1.20 1.46 1.07 1/269 814

T: 0 (  808) P:99 I:200 C:  27902 Min:      5 Act:   13 Avg:   12 Max:   4064
T: 1 (  809) P:99 I:200 C:  27902 Min:      4 Act:   13 Avg:   13 Max:   3838
T: 2 (  810) P:99 I:200 C:  27902 Min:      6 Act:   15 Avg:   14 Max:   3930
```
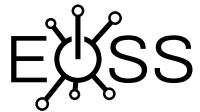
EMBEDDED OPEN SOURCE SUMMIT

OSADL
Open Source Automation Development Lab eG

Jailhouse

- Partitioning hypervisor (based on Linux)
- Licensed under GPL-2.0-only
- https://github.com/siemens/jailhouse
- Supported architectures: x86, ARM (both 32 bit and 64 bit)
- Utilizes the virtualization features of modern CPUs

Jailhouse

Linux

Jailhouse

Linux			

Exclude some memory:
- Either by using "mem=" on the command line or
- by using "reserved-memory" in the devicetree

Jailhouse

Linux			

jailhouse.ko

Jailhouse

Linux (root cell)	Hypervisor (jailhouse.bin)	Communication	Inmate(s)

/dev/jailhouse

Jailhouse

Linux (root cell)	Hypervisor (jailhouse.bin)	Communication	Inmate(s)

EMBEDDED
OPEN SOURCE
SUMMIT

OSADL
Open Source Automation Development Lab eG

Jailhouse

Linux (root cell)	Hypervisor (jailhouse.bin)	Communication	Inmate(s)

Mapped using ioremap() by jailhouse.ko

```
if (ioremap_page_range_sym((unsigned long)vma->addr,
                           (unsigned long)vma->addr + size, phys,
                           PAGE_KERNEL_EXEC)) {
        vunmap(vma->addr);
        return NULL;
    }
```

Jailhouse

Linux (root cell)	Hypervisor (jailhouse.bin)	Communication	Inmate(s)

Mapped using ioremap() by jailhouse.ko

```
if (ioremap_page_range_sym((unsigned long)vma->addr,
                           (unsigned long)vma->addr + size, phys,
                           PAGE_KERNEL_EXEC)) {
        vunmap(vma->addr);
        return NULL;
}
```

EMBEDDED OPEN SOURCE SUMMIT

OSADL
Open Source Automation Development Lab eG

Jailhouse

```
commit 8491502f787c4a902bd4f223b578ef47d3490264
Author: Christoph Hellwig <hch@lst.de>
Date:   Tue Sep 7 19:56:04 2021 -0700

    mm: don't allow executable ioremap mappings

    [...]

diff --git a/mm/vmalloc.c b/mm/vmalloc.c
index e44983fb2d15..3055f04b486b 100644
--- a/mm/vmalloc.c
+++ b/mm/vmalloc.c
@@ -316,7 +316,7 @@ int ioremap_page_range(unsigned long addr, unsigned long end,
 {
        int err;

-       err = vmap_range_noflush(addr, end, phys_addr, prot,
+       err = vmap_range_noflush(addr, end, phys_addr, pgprot_nx(prot),
                                 ioremap_max_page_shift);
        flush_cache_vmap(addr, end);
        return err;
```
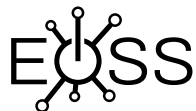
Jailhouse

```
commit 8491502f787c4a902bd4f223b578ef47d3490264
Author: Christoph Hellwig <hch@lst.de>
Date:   Tue Sep 7 19:56:04 2021 -0700

    mm: don't allow executable ioremap mappings

    [...]

diff --git a/mm/vmalloc.c b/mm/vmalloc.c
index e44983fb2d15..3055f04b486b 100644
--- a/mm/vmalloc.c
+++ b/mm/vmalloc.c
@@ -316,7 +316,7 @@ int ioremap_page_range(unsigned long addr, unsigned long end,
 {
        int err;

-       err = vmap_range_noflush(addr, end, phys_addr, prot,
+       err = vmap_range_noflush(addr, end, phys_addr, pgprot_nx(prot),
                                 ioremap_max_page_shift);
        flush_cache_vmap(addr, end);
        return err;
```
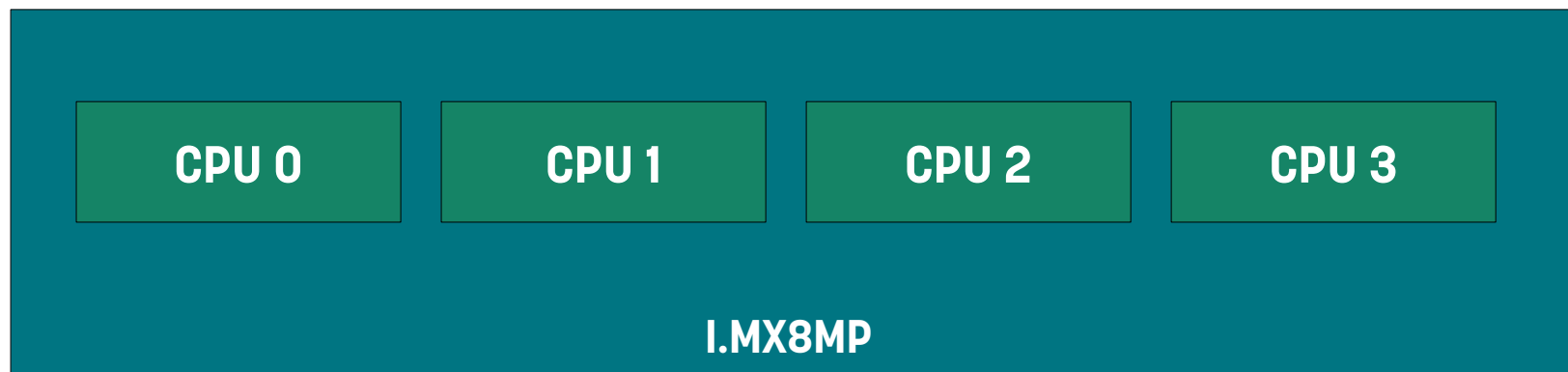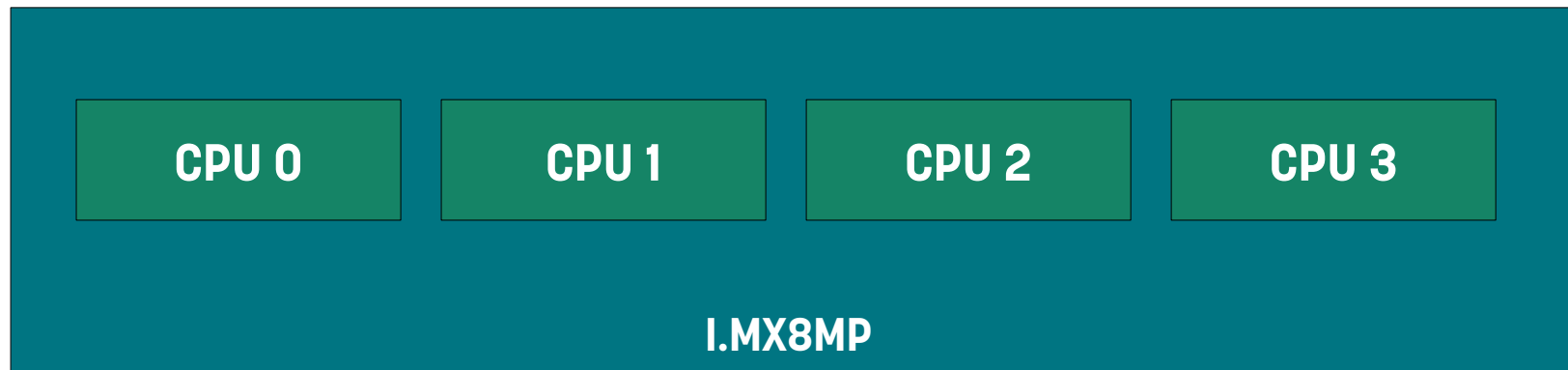
EMBEDDED
OPEN SOURCE
SUMMIT

OSADL
Open Source Automation Development Lab eG

Jailhouse

CPU 0   **CPU 1**   **CPU 2**   **CPU 3**

I.MX8MP

Evaluation of PREEMPT_RT in virtualized environments
Embedded Open Source Summit
June 29, 2023, Prague, Czech Republic

Jailhouse

insmod /root/jailhouse.ko
jailhouse enable configs/arm64/imx8mp.cell

CPU 0	CPU 1	CPU 2	CPU 3

I.MX8MP

Jailhouse

Linux with PREEMPT_RT (root cell)

CPU 0   CPU 1   CPU 2   CPU 3

I.MX8MP

EMBEDDED OPEN SOURCE SUMMIT

Evaluation of PREEMPT_RT in virtualized environments
Embedded Open Source Summit
June 29, 2023, Prague, Czech Republic

OSADL
Open Source Automation Development Lab eG
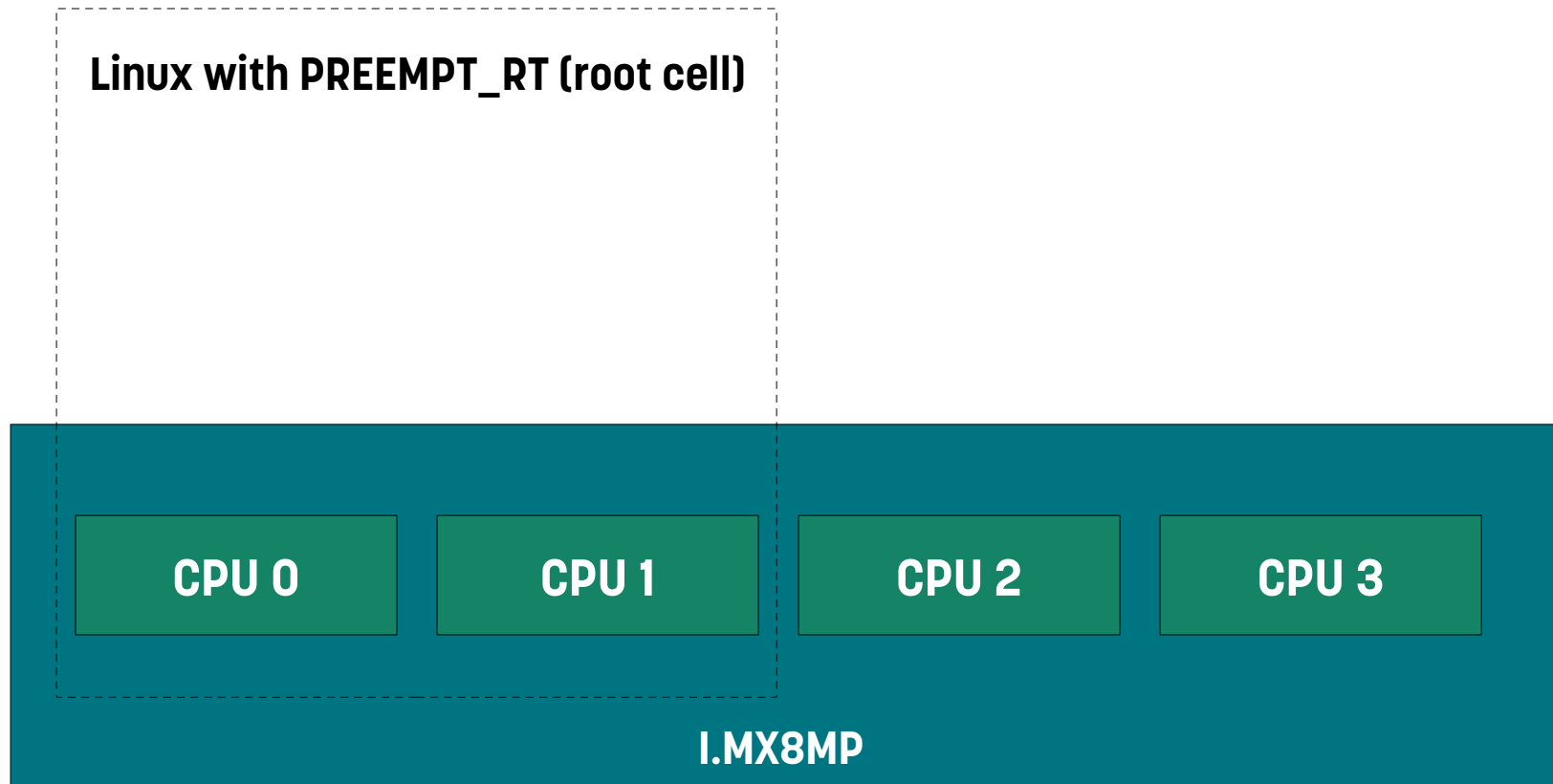
Jailhouse

Linux with PREEMPT_RT (root cell)

```
jailhouse cell linux \
imx8mp-linux-demo.cell \
Image -d \
inmate-imx8mp-pollux.dtb \
-c "clk_ignore_unused \
console=ttymxc0,115200
root=/dev/mmcblk1p2 rootwait rw"
```
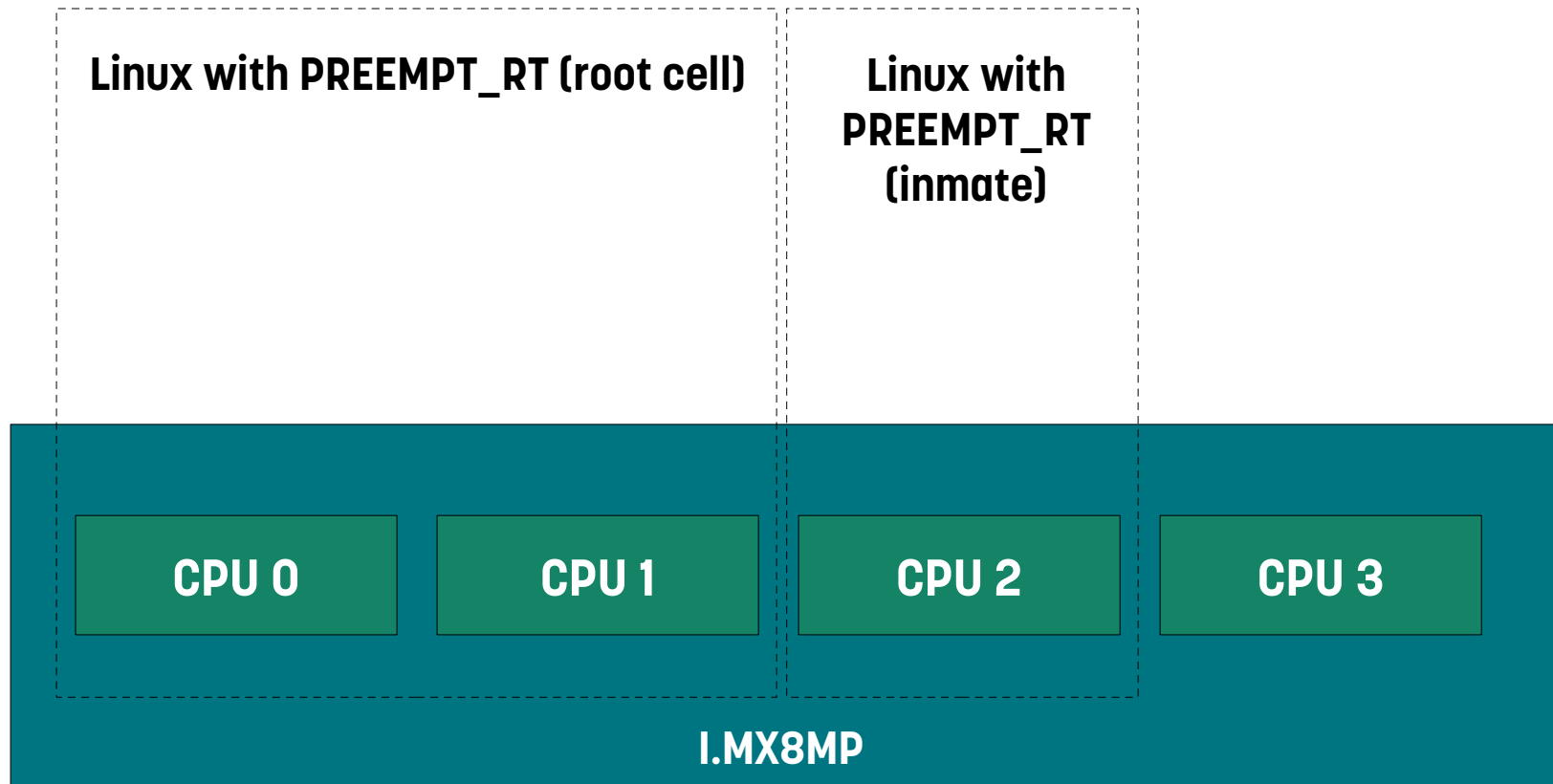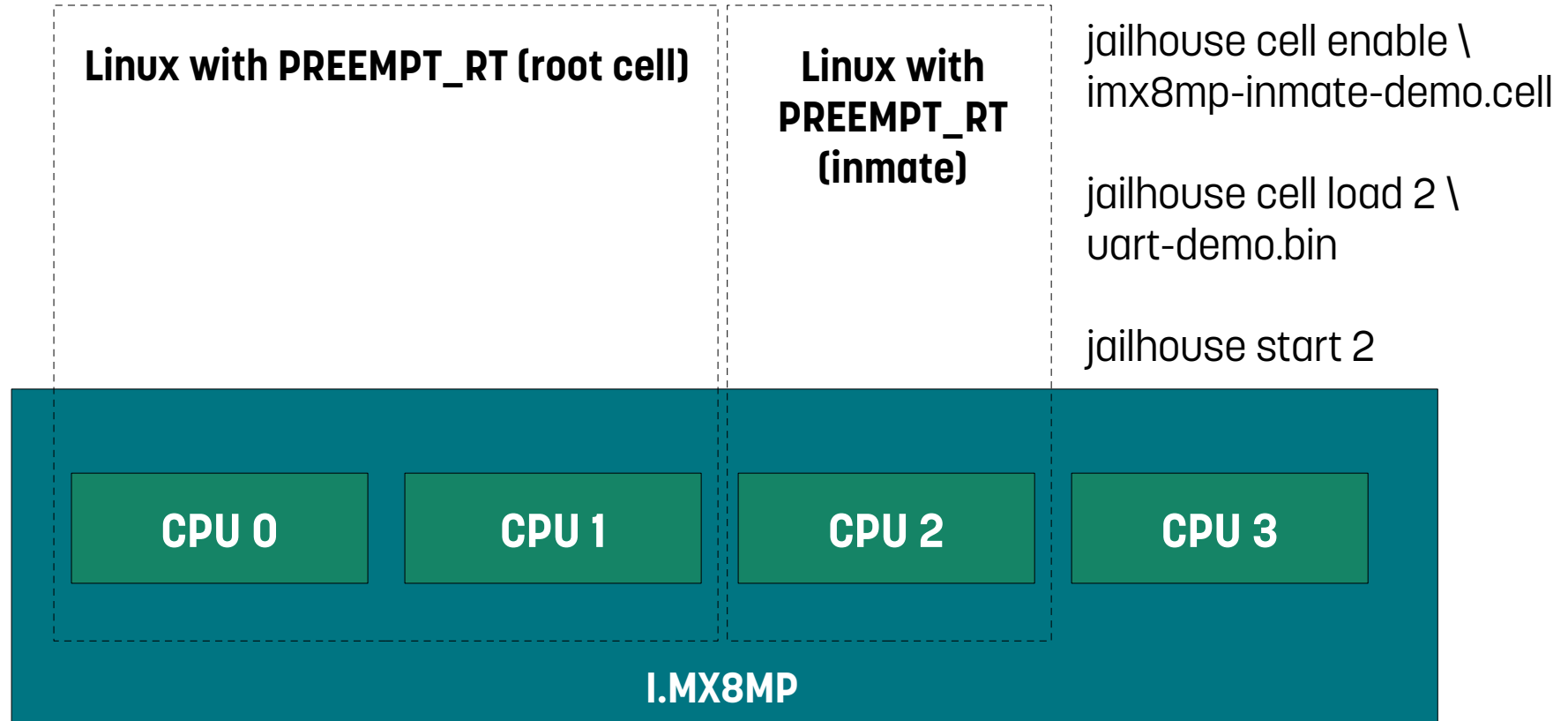
| CPU 0 | CPU 1 | CPU 2 | CPU 3 |

I.MX8MP

Jailhouse

Linux with PREEMPT_RT (root cell)

Linux with PREEMPT_RT (inmate)

CPU 0   CPU 1   CPU 2   CPU 3

I.MX8MP

EMBEDDED OPEN SOURCE SUMMIT

OSADL
Open Source Automation Development Lab eG

Jailhouse

Linux with PREEMPT_RT (root cell)

Linux with PREEMPT_RT (inmate)

jailhouse cell enable \
imx8mp-inmate-demo.cell

jailhouse cell load 2 \
uart-demo.bin

jailhouse start 2

| CPU 0 | CPU 1 | CPU 2 | CPU 3 |

I.MX8MP

Jailhouse

Linux with PREEMPT_RT (root cell)

Linux with PREEMPT_RT (inmate)

bare metal (inmate)

CPU 0 **CPU 1** **CPU 2** **CPU 3**

I.MX8MP

EMBEDDED OPEN SOURCE SUMMIT

OSADL
Open Source Automation Development Lab eG
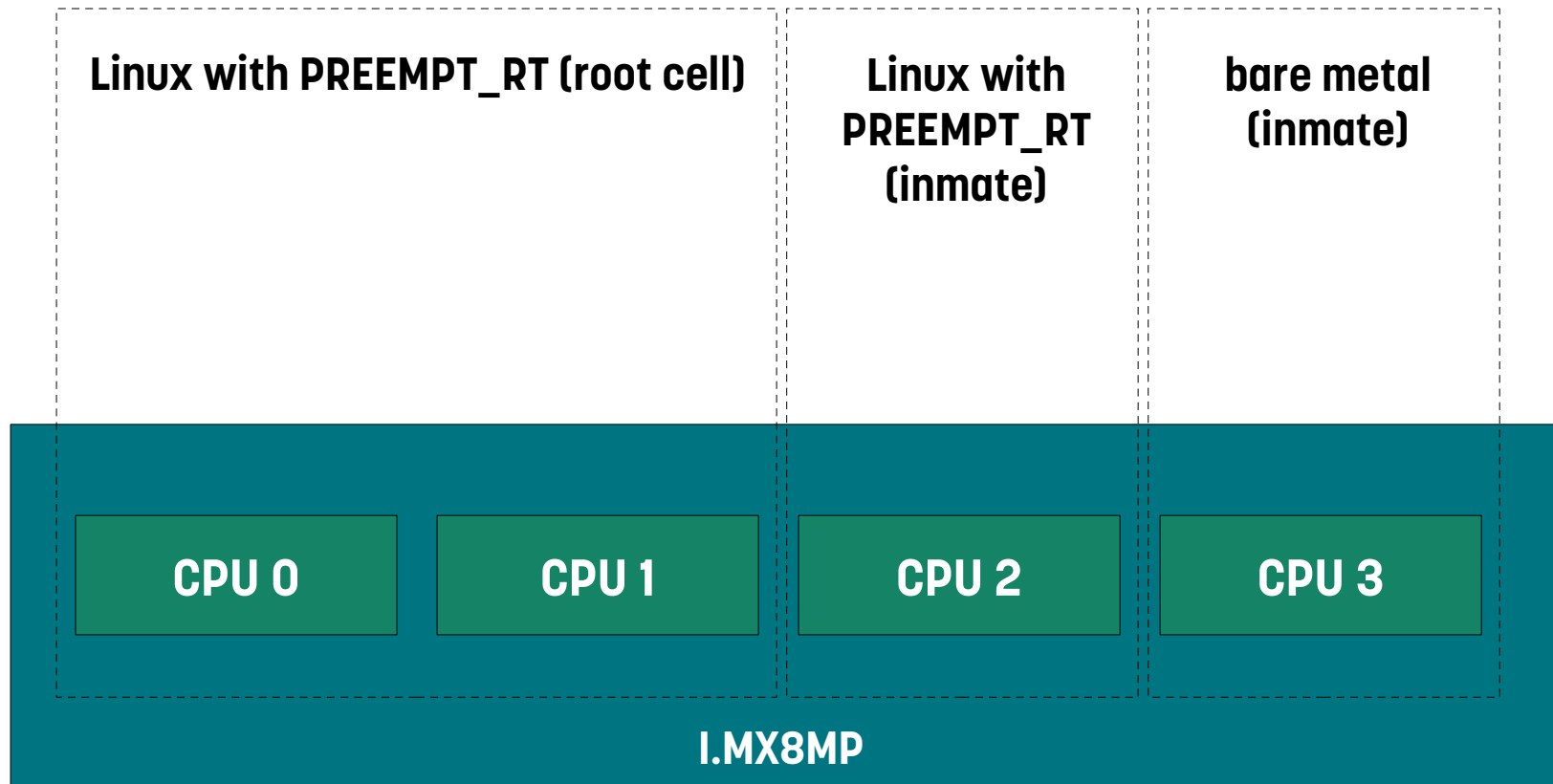
Jailhouse

```
root@jailhouse# jailhouse cell list

ID          Name                    State               Assigned CPUs           Failed CPUs

0           imx8mp                  running             0-1
1           linux-inmate-demo       running             2
2           inmate-demo             running             3
```
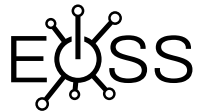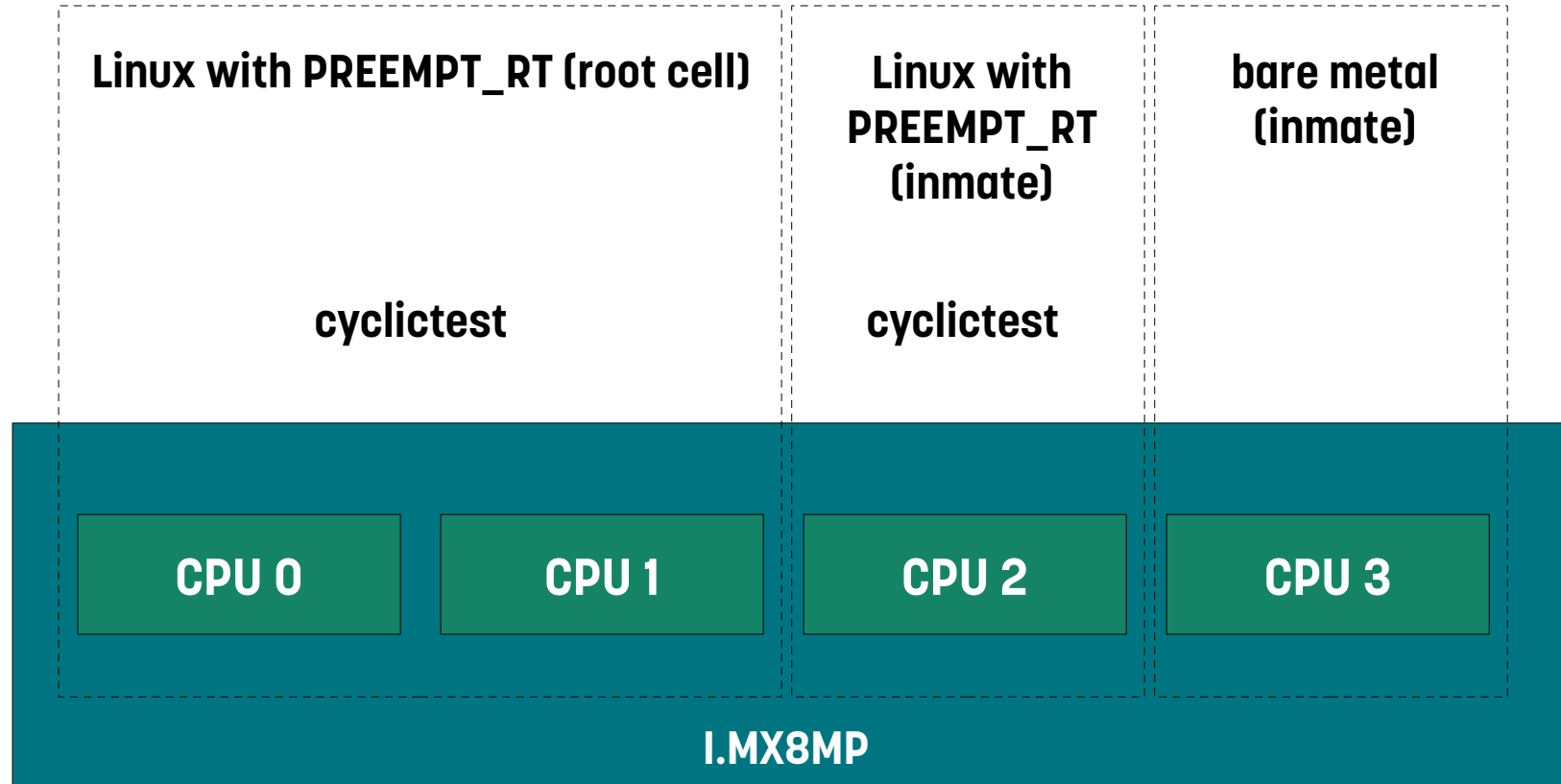
Jailhouse

Linux with PREEMPT_RT (root cell)	Linux with PREEMPT_RT (inmate)	bare metal (inmate)
cyclictest	cyclictest	

| CPU 0 | CPU 1 | CPU 2 | CPU 3 |

I.MX8MP

EMBEDDED OPEN SOURCE SUMMIT

OSADL
Open Source Automation Development Lab eG

Jailhouse: Rootcell

Latency rootcell

Number of latency samples vs Latency (us) - Maximum 45 us (plotted on 15.05.2023 at 14:34:28)

CPU0
CPU1

Evaluation of PREEMPT_RT in virtualized environments
Embedded Open Source Summit
June 29, 2023, Prague, Czech Republic

Jailhouse: Inmate

Latency linuxinmate

CPU0

Number of latency samples

Latency (us) - Maximum 32 us (plotted on 15.05.2023 at 14:34:14)

Evaluation of PREEMPT_RT in virtualized environments
Embedded Open Source Summit
June 29, 2023, Prague, Czech Republic

EOSS
EMBEDDED
OPEN SOURCE
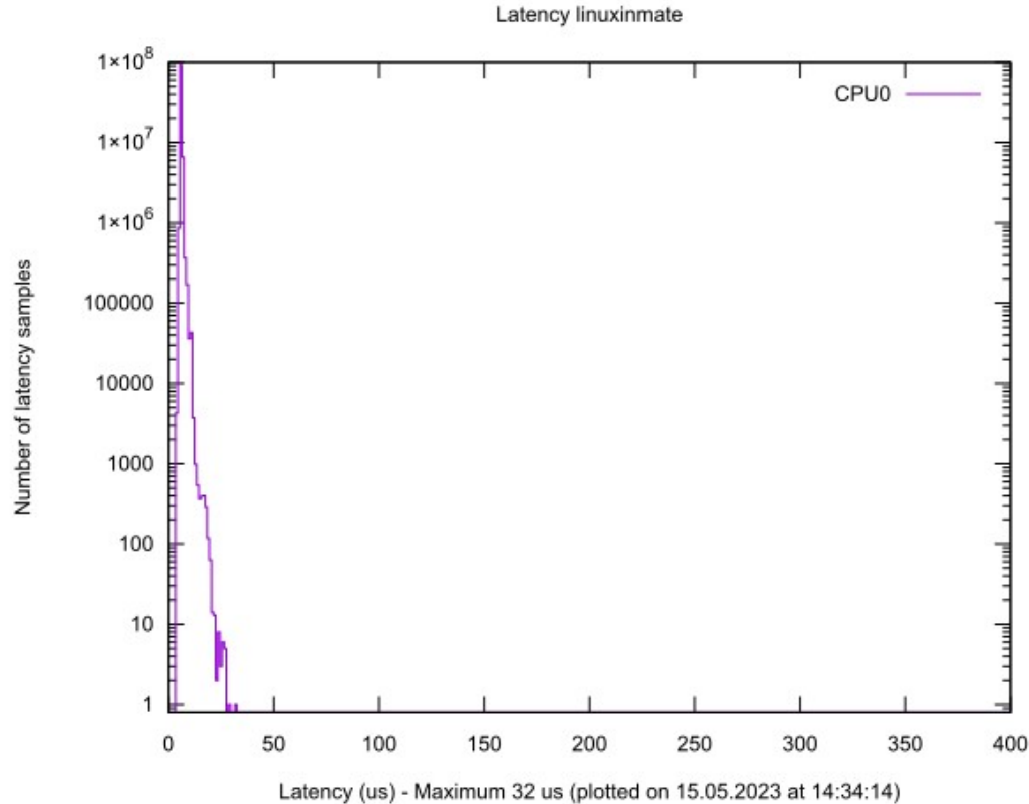SUMMIT

OSADL
Open Source Automation Development Lab eG
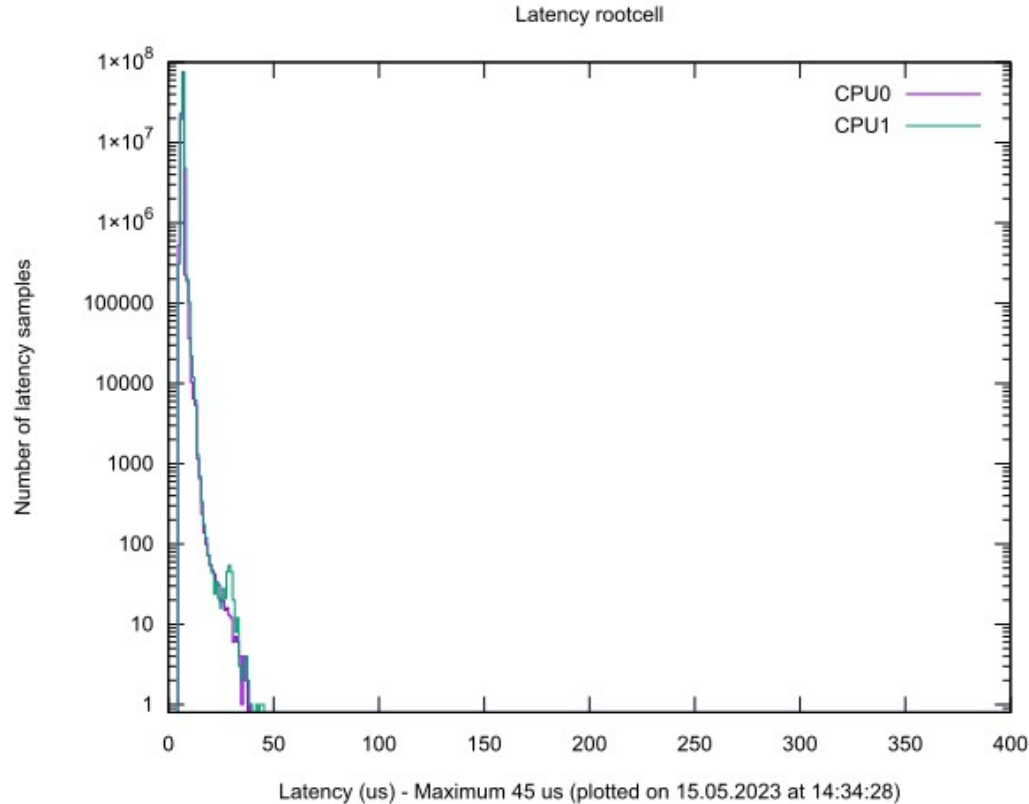
Jailhouse: Blocksys in the root cell

```
root@jailhouse:/# cyclictest -Sp99 -i200 -d0
# /dev/cpu_dma_latency set to 0us
policy: fifo: loadavg: 1.29 0.59 0.24 1/128 446

T: 0 (  445) P:99 I:200 C:  27676 Min:      5 Act:      6 Avg:      6 Max:      30
T: 1 (  446) P:99 I:200 C:   7907 Min:      5 Act:      6 Avg:      6 Max:      17


root@inmate:/# cyclictest -Sp99 -i200 -d0
# /dev/cpu_dma_latency set to 0us
policy: fifo: loadavg: 1.00 0.44 0.17 1/53 109

T: 0 (  109) P:99 I:200 C: 701431 Min:      4 Act:      5 Avg:      5 Max:      22
```

EMBEDDED
OPEN SOURCE
SUMMIT

OSADL
Open Source Automation Development Lab eG

Jailhouse: Blocksys in the root cell

```
root@jailhouse:/# cyclictest -Sp99 -i       -d0
# /dev/cpu_dma_latency set to 0us
policy: fifo: loadavg: 1.29 0.5

                                    blocksys

T: 0 (  445) P:99 I:200 C:  27676 Min:       5 Act:    6 Avg:    6 Max:       30
T: 1 (  446) P:99 I:200 C:   7907 Min:       5 Act:    6 Avg:    6 Max:       17



root@inmate:/# cyclictest -Sp99 -i200 -d0
# /dev/cpu_dma_latency set to 0us
policy: fifo: loadavg: 1.00 0.44 0.17 1/53 109

T: 0 (  109) P:99 I:200 C: 701431 Min:       4 Act:    5 Avg:    5 Max:       22
```

EMBEDDED
OPEN SOURCE
SUMMIT

Evaluation of PREEMPT_RT in virtualized environments
Embedded Open Source Summit
June 29, 2023, Prague, Czech Republic

OSADL
Open Source Automation Development Lab eG

Jailhouse: Blocksys in the root cell

```
root@jailhouse:/# cyclictest -Sp99 -i200 -d0
# /dev/cpu_dma_latency set to 0us
policy: fifo: loadavg: 1.19 0.63 0.27 1/124 455

T: 0 (  448) P:99 I:200 C:  78388 Min:      5 Act:      6 Avg:      6 Max:    5017
T: 1 (  449) P:99 I:200 C:  22385 Min:      5 Act:      9 Avg:      7 Max:    4957


root@inmate:/# cyclictest -Sp99 -i200 -d0
# /dev/cpu_dma_latency set to 0us
policy: fifo: loadavg: 1.00 0.44 0.17 1/53 109

T: 0 (  109) P:99 I:200 C: 882586 Min:      4 Act:      7 Avg:      5 Max:      22
```
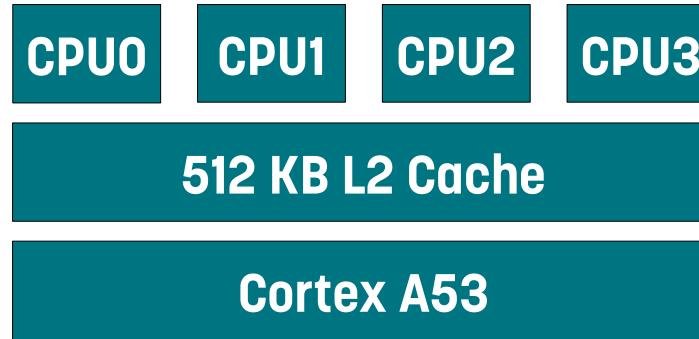
EMBEDDED
OPEN SOURCE
SUMMIT

Evaluation of PREEMPT_RT in virtualized environments
Embedded Open Source Summit
June 29, 2023, Prague, Czech Republic

OSADL
Open Source Automation Development Lab eG

Comparison of technologies
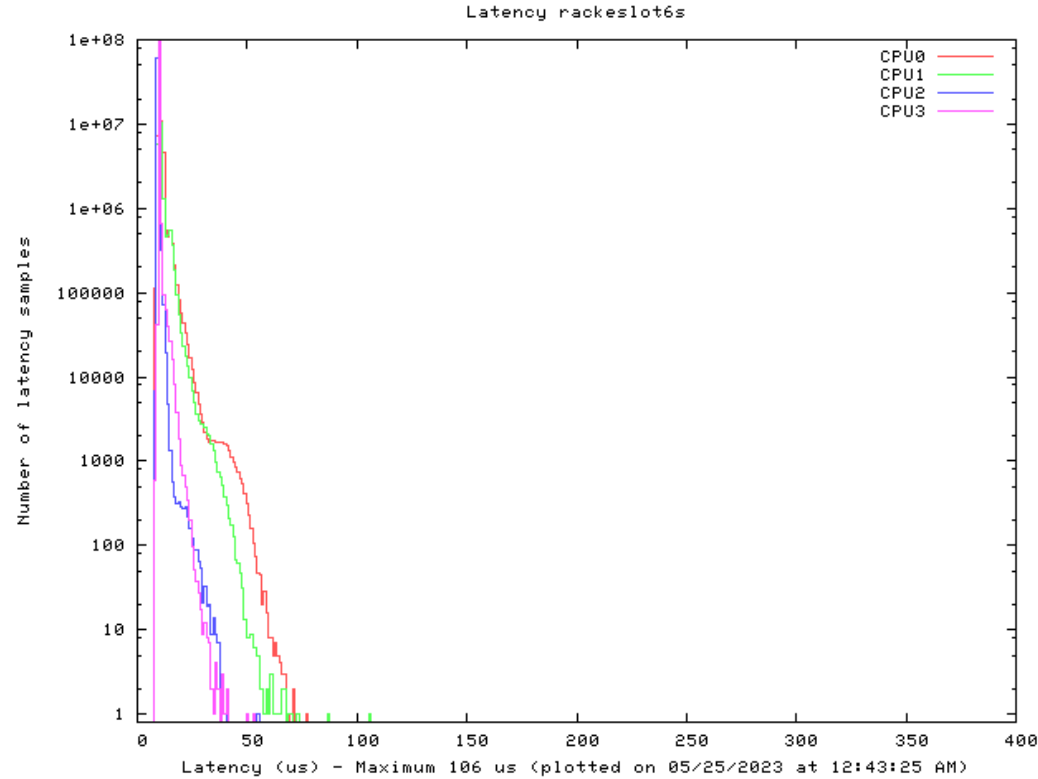
	Docker	KVM	Jailhouse
Real-time in the guest	Yes	Yes (with constraints)	Yes
Real-time with same latency as on host	Yes	Not always	Yes
Separation	Limited	Good	Excellent
Complete independence from host	No	No	No

EMBEDDED OPEN SOURCE SUMMIT

Evaluation of PREEMPT_RT in virtualized environments
Embedded Open Source Summit
June 29, 2023, Prague, Czech Republic

OSADL
Open Source Automation Development Lab eG

Comparison of technologies

	Docker	KVM	Jailhouse
Real-time in the guest	Yes	Yes (with constraints)	Yes
Real-time with same latency as on host	Yes	Not always	Yes
Separation	Limited	Good	Excellent
Complete independence from host	No	No	No

But why?!

Evaluation of PREEMPT_RT in virtualized environments
Embedded Open Source Summit
June 29, 2023, Prague, Czech Republic

OSADL
Open Source Automation Development Lab eG

Shared last level caches

CPU0 **CPU1** **CPU2** **CPU3**

512 KB L2 Cache

Cortex A53

Shared last level caches

Evaluation of PREEMPT_RT in virtualized environments
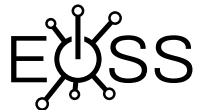Embedded Open Source Summit
June 29, 2023, Prague, Czech Republic

Shared last level caches

```
# stress-ng --malloc 1 --malloc-bytes 512k --malloc-max 32 --taskset 0

# cyclictest -n -m -Sp98 -i3000 -h400
policy: fifo: loadavg: 3.67 3.51 3.47 2/146 750

T: 0 (  395) P:98 I:3000 C:4942427 Min:      6 Act:   50 Avg:   46 Max:      104
T: 1 (  396) P:98 I:3000 C:4942422 Min:      6 Act:   26 Avg:   69 Max:      623
T: 2 (  397) P:98 I:3000 C:4942417 Min:      6 Act:   51 Avg:   57 Max:      490
T: 3 (  398) P:98 I:3000 C:4942413 Min:      6 Act:   39 Avg:   76 Max:      313
```

EMBEDDED
OPEN SOURCE
SUMMIT

OSADL
Open Source Automation Development Lab eG

Shared last level caches

```
# stress-ng --malloc 1 --malloc-bytes 512k --malloc-max 32 --taskset 0

# cyclictest -n -m -Sp98 -i3000 -h400
policy: fifo: loadavg: 3.67 3.51 3.47 2/146 750


T: 0 (  395) P:98 I:3000 C:4942427 Min:      6 Act:    50 Avg:    46 Max:     104
T: 1 (  396) P:98 I:3000 C:4942422 Min:      6 Act:    26 Avg:    69 Max:     623
T: 2 (  397) P:98 I:3000 C:4942417 Min:      6 Act:    51 Avg:    57 Max:     490
T: 3 (  398) P:98 I:3000 C:4942413 Min:      6 Act:    39 Avg:    76 Max:     313
```
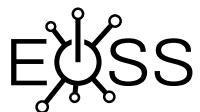
CPU0	CPU1	CPU2	CPU3

512 KB L2 Cache

Cortex A53

Shared last level caches

```
# stress-ng --malloc 1 --malloc-bytes 512k --malloc-max 32 --taskset 0

# cyclictest -n -m -Sp98 -i3000 -h400
policy: fifo: loadavg: 3.67 3.51 3.47 2/146 750


T: 0 (  395) P:98 I:3000 C:4942427 Min:        6 Act:   50 Avg:   46 Max:    104
T: 1 (  396) P:98 I:3000 C:4942422 Min:        6 Act:   26 Avg:   69 Max:    623
T: 2 (  397) P:98 I:3000 C:4942417 Min:        6 Act:   51 Avg:   57 Max:    490
T: 3 (  398) P:98 I:3000 C:4942413 Min:        6 Act:   39 Avg:   76 Max:    313
```
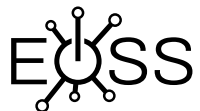
| CPU0 | CPU1 | CPU2 | CPU3 |

512 KB L2 Cache

Cortex A53

Shared last level caches

```
# stress-ng --malloc 1 --malloc-bytes 512k --malloc-max 32 --taskset 0

# cyclictest -n -m -Sp98 -i3000 -h400
policy: fifo: loadavg: 3.67 3.51 3.47 2/146 750


T: 0 (  395) P:98 I:3000 C:4942427 Min:      6 Act:   50 Avg:   46 Max:    104
T: 1 (  396) P:98 I:3000 C:4942422 Min:      6 Act:   26 Avg:   69 Max:    623
T: 2 (  397) P:98 I:3000 C:4942417 Min:      6 Act:   51 Avg:   57 Max:    490
T: 3 (  398) P:98 I:3000 C:4942413 Min:      6 Act:   39 Avg:   76 Max:    313
```

| CPU0 | CPU1 | CPU2 | CPU3 |

512 KB L2 Cache

Cortex A53

EMBEDDED OPEN SOURCE SUMMIT

OSADL
Open Source Automation Development Lab eG

Shared last level caches

| CPU0 | CPU1 | CPU2 | CPU3 | | CPU4 | CPU5 |

| 1 MB L2 Cache |
| 4x Cortex A53 |

| 1 MB L2 Cache |
| 2x Cortex A73 |

Shared last level caches

```
# stress-ng --malloc 2 --malloc-bytes 512k --malloc-max 32 --taskset 0-1

# cyclictest -n -m -Sp98 -i3000 -h400
policy: fifo: loadavg: 2.90 1.21 0.50 3/240 4476

T: 0 ( 4466) P:98 I:3000 C:    7352 Min:       6 Act:     38 Avg:     30 Max:       389
T: 1 ( 4467) P:98 I:3000 C:    7348 Min:       7 Act:     30 Avg:     40 Max:       415
T: 2 ( 4468) P:98 I:3000 C:    7343 Min:       6 Act:     19 Avg:     23 Max:       373
T: 3 ( 4469) P:98 I:3000 C:    7338 Min:       6 Act:     12 Avg:     15 Max:       315
T: 4 ( 4470) P:98 I:3000 C:    7334 Min:       4 Act:      5 Avg:      6 Max:        51
T: 5 ( 4471) P:98 I:3000 C:    7329 Min:       4 Act:      5 Avg:      7 Max:        68
```
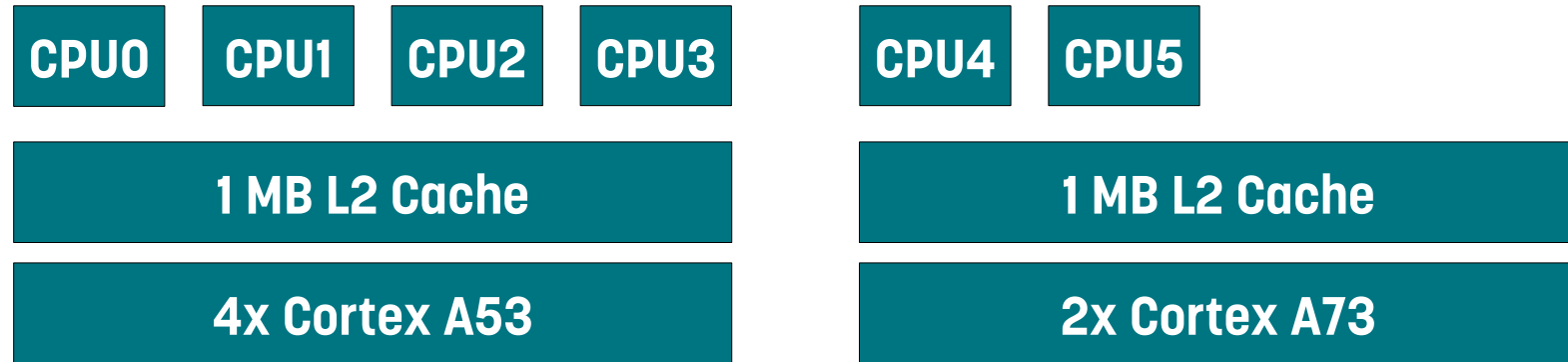
EMBEDDED
OPEN SOURCE
SUMMIT

Evaluation of PREEMPT_RT in virtualized environments
Embedded Open Source Summit
June 29, 2023, Prague, Czech Republic

OSADL
Open Source Automation Development Lab eG

Shared last level caches

```
# stress-ng --malloc 2 --malloc-bytes 512k --malloc-max 32 --taskset 0-1

# cyclictest -n -m -Sp98 -i3000 -h400
policy: fifo: loadavg: 2.90 1.21 0.50 3/240 4476

T: 0 ( 4466) P:98 I:3000 C:     7352 Min:      6 Act:    38 Avg:    30 Max:     389
T: 1 ( 4467) P:98 I:3000 C:     7348 Min:      7 Act:    30 Avg:    40 Max:     415
T: 2 ( 4468) P:98 I:3000 C:     7343 Min:      6 Act:    19 Avg:    23 Max:     373
T: 3 ( 4469) P:98 I:3000 C:     7338 Min:      6 Act:    12 Avg:    15 Max:     315
```
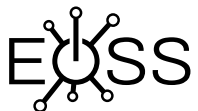
CPU0		CPU1		CPU2		CPU3		CPU4		CPU5				5 Avg:	6 Max:	51
														5 Avg:	7 Max:	68

```
          4          00     7334                              5 Avg:     6 Max:      51
          4          00     7329                              5 Avg:     7 Max:      68
```

1 MB L2 Cache	1 MB L2 Cache
4x Cortex A53	2x Cortex A73

EMBEDDED OPEN SOURCE SUMMIT

Evaluation of PREEMPT_RT in virtualized environments
Embedded Open Source Summit
June 29, 2023, Prague, Czech Republic

OSADL
Open Source Automation Development Lab eG

Shared last level caches

```
# stress-ng --malloc 2 --malloc-bytes 512k --malloc-max 32 --taskset 0-1

# cyclictest -n -m -Sp98 -i3000 -h400
policy: fifo: loadavg: 2.90 1.21 0.50 3/240 4476


T: 0 ( 4466) P:98 I:3000 C:      7352 Min:       6 Act:    38 Avg:    30 Max:     389
T: 1 ( 4467) P:98 I:3000 C:      7348 Min:       7 Act:    30 Avg:    40 Max:     415
T: 2 ( 4468) P:98 I:3000 C:      7343 Min:       6 Act:    19 Avg:    23 Max:     373
T: 3 ( 4469) P:98 I:3000 C:      7338 Min:       6 Act:    12 Avg:    15 Max:     315
```

| CPU0 | CPU1 | CPU2 | CPU3 | | CPU4 | CPU5 |

```
7334              5 Avg:     6 Max:      51
7329              5 Avg:     7 Max:      68
```

| 1 MB L2 Cache | 1 MB L2 Cache |

| 4x Cortex A53 | 2x Cortex A73 |

EMBEDDED
OPEN SOURCE
SUMMIT

Evaluation of PREEMPT_RT in virtualized environments
Embedded Open Source Summit
June 29, 2023, Prague, Czech Republic

OSADL
Open Source Automation Development Lab eG

Shared last level caches

```
# stress-ng --malloc 2 --malloc-bytes 512k --malloc-max 32 --taskset 0-1

# cyclictest -n -m -Sp98 -i3000 -h400
policy: fifo: loadavg: 2.90 1.21 0.50 3/240 4476
```

```
T: 0 ( 4466) P:98 I:3000 C:    7352 Min:      6 Act:   38 Avg:   30 Max:      389
T: 1 ( 4467) P:98 I:3000 C:    7348 Min:      7 Act:   30 Avg:   40 Max:      415
T: 2 ( 4468) P:98 I:3000 C:    7343 Min:      6 Act:   19 Avg:   23 Max:      373
T: 3 ( 4469) P:98 I:3000 C:    7338 Min:      6 Act:   12 Avg:   15 Max:      315
```


| CPU0 | CPU1 | CPU2 | CPU3 | | CPU4 | CPU5 |

| 1 MB L2 Cache | 1 MB L2 Cache |

| 4x Cortex A53 | 2x Cortex A73 |

EMBEDDED
OPEN SOURCE
SUMMIT

OSADL
Open Source Automation Development Lab eG

Summary with respect to real-time capabilities

- A guest system may have real-time capabilities.
- Containers (e.g. Docker) can have the same latencies as the host.
- Full hardware virtualization guest systems with KVM may have longer latencies than the host.
- With Jailhouse full hardware virtualization guest systems can have the same latencies as the host.
- But in any case, due to shared hardware resources, the host system can affect the guest's real-time and the other way round.

EMBEDDED OPEN SOURCE SUMMIT

Evaluation of PREEMPT_RT in virtualized environments
Embedded Open Source Summit
June 29, 2023, Prague, Czech Republic

OSADL
Open Source Automation Development Lab eG

Summary with respect to separation

- A guest system may or may not be separated from the host system.

- Containers are normally not well separated from each other or from the host.

- Full hardware virtualization guest systems with KVM may provide separation to a high degree.

- With Jailhouse host and guest systems may well be separated from each other.

EMBEDDED
OPEN SOURCE
SUMMIT

Evaluation of PREEMPT_RT in virtualized environments
Embedded Open Source Summit
June 29, 2023, Prague, Czech Republic

OSADL
Open Source Automation Development Lab eG